



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Programming and reprogramming neural cell types using synthetic transcription factors

Mantas Matjusaitis

Supervisor: Prof Steven M. Pollard, PhD

MRC Centre for Regenerative Medicine,

The University of Edinburgh,

2017

A thesis submitted to the University of Edinburgh for the degree of Doctor of
Philosophy



Declaration of Originality

I hereby declare that this thesis and the work presented in it was composed by and originated entirely from me and has not been submitted in any form for another degree or diploma at any other institution. Any information derived from the published or unpublished work of others has been acknowledged in the text and references are listed in the bibliography.

Date

Signature

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivations license. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.

Table of contents

DECLARATION OF ORIGINALITY	2
COPYRIGHT DECLARATION	2
TABLE OF CONTENTS	3
ABSTRACT	10
LAY SUMMARY	12
ACKNOWLEDGEMENTS	14
LIST OF ABBREVIATIONS	15
LIST OF FIGURES	17
LIST OF TABLES	21
1 INTRODUCTION	22
1.1 OVERVIEW	22
1.2 PROGRAMING AND REPROGRAMING CELL TYPES	24
1.2.1 CELL DIFFERENTIATION DURING DEVELOPMENT AND <i>IN VITRO</i>	26
1.2.2 REPROGRAMING REVERSES DIFFERENTIATION TO GENERATE PLURIPOTENT STEM CELLS	29
1.2.3 TRANSDIFFERENTIATION: DIRECT REPROGRAMING OF SOMATIC CELLS	31
1.2.4 MASTER TRANSCRIPTION FACTORS AND THEIR MODE OF ACTION.	33
1.2.5 NEW WAYS OF INDUCING CELL TYPE CHANGES.	36
1.3 GENOME EDITING TECHNOLOGIES AND SYNTHETIC TRANSCRIPTION FACTORS	37
1.3.1 INTRODUCTION TO GENOME EDITING TOOLS – ZINC FINGER NUCLEASES, TALENS AND CRISPR-Cas9	37

1.3.2 BIOLOGY OF CRISPR-Cas9	40
1.3.3 SYNTHETIC TRANSCRIPTION FACTORS	42
1.3.4 EFFECTORS USED WITH SYNTHETIC TRANSCRIPTION FACTORS	47
1.3.5 CORE PROMOTERS AND TRANSCRIPTION MACHINERY IN EUKARYOTES	51
1.3.6 SYNTHETIC TRANSCRIPTION FACTORS IN CELL TYPE PROGRAMING	55
1.4 NEURAL STEM CELLS AND OLIGODENDROCYTE DIFFERENTIATION	58
1.4.1 NEURAL STEM CELLS CAN DIFFERENTIATE INTO OLIGODENDROCYTES, ASTROCYTES AND NEURONS	58
1.4.2 GENETIC PROFILE OF OLIGODENDROCYTES	59
1.4.3 METHODS FOR GENERATION OF OLIGODENDROCYTES EX VIVO	62
1.5 AIMS AND HYPOTHESIS	65
<u>2 METHODS AND MATERIALS</u>	<u>66</u>
2.1 CELL CULTURE	66
2.1.1 NEURAL STEM CELL CULTURE MEDIUM	66
2.1.2 MOUSE EMBRYONIC FIBROBLAST CULTURES	66
2.1.3 MEF CULTURE MEDIUM	66
2.1.4 HEK293T CULTURE MEDIUM	67
2.1.5 FREEZING AND THAWING CELL LINES	67
2.1.6 NSCs DIFFERENTIATION MEDIUM AND PROTOCOL	67
2.1.7 MEF TRANSDIFFERENTIATION TO OPCs	68
2.1.8 CLONAL CELL LINE DERIVATION	69
2.2 MAMMALIAN CELL TRANSFECTION (MEF, NS AND HEK293T)	69
2.2.1 POLYETHYLENIMINE (PEI)	69
2.2.2 NUCLEOFECTION	69
2.2.3 LIPOFECTION	70

2.3	IMMUNOCYTOCHEMISTRY	70
2.4	MICROSCOPY AND IMAGE PROCESSING	71
2.5	IMAGE CYTOMETRY	71
2.6	FLOW CYTOMETRY	72
2.7	TAQMAN QUANTITATIVE RT-PCR	72
2.7.1	RNA EXTRACTION	72
2.7.2	REVERSE TRANSCRIPTASE (RT) REACTION	73
2.7.3	QPCR	73
2.8	RNA-FISH	74
2.9	MOLECULAR CLONING	74
2.9.1	SITE-DIRECTED MUTAGENESIS	74
2.9.2	GATEWAY CLONING	75
2.9.3	GIBSON CLONING	75
2.9.4	GOLDEN GATE ASSEMBLY	75
2.9.5	<i>E. COLI</i> TRANSFORMATION	76
2.10	DESIGN AND PRODUCTION OF U6-GRNA EXPRESSION PLASMIDS	78
2.11	GENE TAGGING WITH V5 USING SINGLE-STRAND OLIGONUCLEOTIDE	78
3	<u>REGULATING MOUSE GENE EXPRESSION USING STFS</u>	80
3.1	ACTIVATING <i>SOX10</i>, <i>OLIG2</i> AND <i>NKX6-2</i> IN MOUSE NS CELLS AND MEFS	80
3.1.1	CONSTRUCTION OF U6-DRIVEN GRNA EXPRESSION VECTORS	80
3.1.2	<i>SOX10</i> , <i>OLIG2</i> AND <i>NKX6-2</i> ACTIVATION IN MOUSE NS CELLS AND MEFS	83
3.2	EXPLORING THE SYNERGISTIC VERSUS ADDITIVE EFFECTS OF MULTIPLE GRNAs USING SINGLE CELL ASSAYS	88
3.2.1	OPTIMISING A QUANTITATIVE IMMUNOCYTOCHEMISTRY ('IMAGE CYTOMETRY') ASSAY	88

3.2.2	TRANSCRIPTIONAL ACTIVATION OF <i>OLIG2</i> IN HEK293T USING MULTIPLE GRNAS	91
3.2.3	QUANTIFICATION OF ABSOLUTE MRNA LEVELS WITHIN SINGLE CELLS USING RNA-FISH	95
3.2.4	dCas9 AND <i>Sox10</i> MRNAS MOLECULES DO NOT ALWAYS CO-LOCALIZE IN SINGLE MEF CELLS	95
3.3	<i>OLIG2</i> ACTIVATION IN DIFFERENTIATING NEURAL STEM CELLS	99
3.3.1	CHARACTERISATION OF A PDGFRA-GFP (PG1.1) CELL LINE	99
3.3.2	<i>OLIG2</i> ACTIVATION DURING PG1.1 DIFFERENTIATION TO OPCs OR ASTROCYTES	101
3.4	INVESTIGATING POTENTIAL OF DIFFERENT TRANSCRIPTION ACTIVATING DOMAINS (TADs) TO OVERCOME BIOLOGICAL RESTRICTIONS FOR ACTIVATION	105
3.4.1	SCREENING OF NEW TRANSCRIPTION ACTIVATING DOMAINS (TADs)	105
3.5	MOUSE GENE REPRESSION IN NSC USING dCas9-KRAB	108
3.5.1	GRNA PLACEMENT FOR GENE REPRESSION	108
3.5.2	<i>Sox9</i> REPRESSION IN SOX9-V5 NS CELL LINE	108
3.6	CONCLUDING REMARKS	112
4	<u>MULTIPLEX DELIVERY OF STFS: A NEW TOOLKIT FOR PLASMID ASSEMBLY</u>	114
4.1	DESIGN AND VALIDATION OF AN ALL-IN ONE PLASMID FOR MULTIPLEX TRANSCRIPTIONAL ACTIVATION	114
4.1.1	DESIGNING AND VALIDATING A NOVEL EXTENSIBLE MAMMALIAN MODULAR EXPRESSION PLASMID ASSEMBLY (EMMA) SYSTEM	114
4.1.2	CONSTRUCTION OF TOOLKIT FOR GENE REGULATION IN MULTIPLEX	120
4.2	CONSTRUCTION AND VALIDATION OF SYSTEM FOR SIMULTANEOUS REPRESSION AND ACTIVATION	132
4.2.1	CONSTRUCTION AND VALIDATION OF VRER dCas9 MUTANTS FOR ALTERED PAM SPECIFICITY	134
4.2.2	MOUSE GENE REPRESSION IN NSC USING VRER-dCas9-KRAB	138
4.3	CONSTRUCTION OF A <i>Sox10</i>-V5 TAGGED NSC LINE FOR MULTIPLEXING VALIDATION	141

4.3.1	EVALUATING SOX10 ANTIBODY SPECIFICITY IN OUR CELL LINES	141
4.3.2	TAGGING <i>Sox10</i> WITH C-TERMINUS V5 IN PG1.1 AND BL6 NSC LINES	141
4.4	CONCLUDING REMARKS	148
5	<u>PROGRAMING AND REPROGRAMING CELL IDENTITY USING STFS</u>	150
5.1	DIRECTING NS CELL DIFFERENTIATION BY OVEREXPRESSION OR STF-DRIVEN ACTIVATION OF KEY OPC MASTER REGULATORS	150
5.1.1	CONSTRUCTION OF PG1.1 CELL LINE WITH INDUCIBLE SOX10 AND OLIG2 OVEREXPRESSION (PG1.1-SO)	150
5.1.2	CONSTRUCTION OF PG1.1 CELL LINE FOR <i>Sox10</i> ACTIVATION WITH dCas9-VP160 (PG1.1- 3SON)	156
5.1.3	EFFECT OF <i>SOX10</i> AND <i>OLIG2</i> CDNA OVEREXPRESSION AND <i>Sox10</i> ACTIVATION WITH STFS ON MOUSE NEURAL STEM CELLS IN SELF-RENEWAL CONDITIONS	158
5.1.4	EFFECT OF <i>SOX10</i> AND <i>OLIG2</i> CDNA OVEREXPRESSION AND <i>Sox10</i> TRANSCRIPTION ACTIVATION (STFS) ON MOUSE NEURAL STEM CELLS IN OPC DIFFERENTIATION CONDITIONS	161
5.1.5	EFFECT OF <i>Sox10</i> TRANSCRIPTION ACTIVATION BY STFS ON BL6 NSC DIFFERENTIATION TOWARDS OL LINEAGE	165
5.2	EFFECT OF <i>Sox10</i> , <i>OLIG2</i> AND <i>Nkx6-2</i> TRANSCRIPTION ACTIVATION ON MEF TRANSDIFFERENTIATION TO OL	170
5.2.1	GENERATION OF MEF LINES FROM MOUSE EMBRYOS	170
5.2.2	MEF TRANSDIFFERENTIATION TO OPC/OLS AFTER <i>Sox10</i> , <i>OLIG2</i> AND <i>Nkx6-2</i> ACTIVATION WITH STFS (3SON)	172
5.3	CONCLUDING REMARKS	177
6	<u>GENERAL DISCUSSION</u>	178
6.1	UNDERSTANDING dCas9 SYNTHETIC TRANSCRIPTION FACTORS	178

6.1.1	STFS HAVE CELL-TYPE SPECIFIC FUNCTIONAL CONSTRAINTS THAT LIMITS PREDICTIVE DESIGN RULES	178
6.1.2	DELIVERY OF MULTIPLE GRNAs INCREASES THE SUCCESS OF TRANSCRIPTIONAL ACTIVATION	182
6.1.3	DCas9-VP160 AMOUNT DOES NOT SEEM TO BE A KEY LIMITING FACTOR	186
6.1.4	STRONGER TADs ALLOW HIGHER GENE ACTIVATION BUT ARE STILL RESTRICTED BY BIOLOGICAL BARRIERS	186
6.1.5	EPIGENETIC MODIFICATIONS AS A MAIN LIMITING FACTOR FOR EFFICIENT AND ROBUST STF FUNCTION.	188
6.2	TOOLKIT FOR SIMULTANEOUS GENE ACTIVATION AND REPRESSION IN MULTIPLEX	189
6.2.1	AN EXTENSIBLE MAMMALIAN MODULAR ASSEMBLY (EMMA), SYSTEM FOR SIMPLE AND RAPID CONSTRUCTION OF EXPRESSION PLASMIDS	189
6.2.2	SIMULTANEOUS ACTIVATION OF MULTIPLE ENDOGENOUS GENES BY DRIVING EXPRESSION OF THREE GRNAs-tRNA UNITS FROM SINGLE U6 PROMOTER	190
6.2.3	ALL-IN ONE PLASMID, DRIVEN BY MULTIPLE U6 PROMOTERS, FOR SIMULTANEOUS ACTIVATION OF MULTIPLE ENDOGENOUS GENES	192
6.2.4	VRER MUTANT DCas9 RECOGNIZES NGCG INSTEAD OF NGG PAM	193
6.3	PROGRAMING AND REPROGRAMING NSC AND MEFs USING SYNTHETIC TRANSCRIPTION FACTORS	196
6.3.1	SOX10 OVEREXPRESSION OR ACTIVATION DRIVES NSC DIFFERENTIATION TOWARDS OL LINEAGE	197
6.3.2	<i>Sox10</i> , <i>Olig2</i> AND <i>Nkx6-2</i> ACTIVATION BY STFS DRIVE MEF TRANSDIFFERENTIATION TOWARDS OL LINEAGE	199
6.4	FUTURE WORK	200
7	REFERENCES	203
8	APPENDIX	233

8.1	SEQUENCES OF GRNAs USED IN THIS STUDY	233
8.2	LIST OF CUSTOM-MADE RNA-FISH PROBES AND THEIR SEQUENCES	235
8.3	LIST OF KEY PCR PRIMER USED IN THIS STUDY	238

Abstract

Production of large numbers of desirable human cell types in the laboratory is one of the major goals of stem cell research. Current experimental approaches have focused on the strategy of recapitulating the events of normal embryogenesis in culture, by treating cells – either tissue stem cells or pluripotent stem cells (iPS/ES cells) – with cocktails of growth factors, matrix proteins or pharmacological agents. This is challenging and often requires weeks or months of elaborate cell culture regimes. An alternative approach is the forced expression of master regulatory transcription factors; this can bypass developmental programs and drive conversion to the target cell type. Each of these strategies is inefficient and unreliable. Recently a new opportunity has arisen to exploit synthetic transcription factors (sTFs) to program and reprogram cell fate. To create such sTFs the CRISPR/Cas9 system is repurposed through tethering of catalytically dead Cas9 to various transcriptional regulatory effector domains (e.g. VP16, KRAB).

In this thesis, we have explored sTFs as tools to reset transcriptional regulatory networks in neural stem cells and mouse embryonic fibroblasts. We tested transcriptional activation of key neural lineage target genes (e.g. *Olig2*, *Sox10* and *Nkx6.2*). We designed and validated a series of sTFs that could effectively activate these. We have found that activation of *Sox10* by dCas9-VP160 in mouse neural stem cells can increase the amount of arising oligodendrocyte and oligodendrocyte precursor cells during the differentiation.

The activity of sTFs strongly depends on cellular context: i.e. a specific sTF might work well in one cell type but not another. Importantly, these biological barriers

are not easily overcome by increasing the strength of the sTF – either through levels or types of effector domains used. Our data inspecting single cells suggests that multiplex delivery of sTFs can indeed cooperate by both increasing the number of cells that activated the gene of interest and increasing the level of transcriptional activation in a given cell. To fully exploit these new technologies, we therefore developed a new construction pipeline that allows easy and efficient assembly of multiple sTFs. Using this approach, we were able to successfully activate three different target genes from a single expression plasmid (*Olig2*, *Sox10* and *Nkx6.2*) in fibroblasts. These sTFs were able to force fibroblast transdifferentiation towards oligodendrocyte lineage. Future studies will explore further how to exploit these sTFs to augment or replace current reprogramming strategies.

Lay summary

Many human diseases, such as multiple sclerosis and Parkinson's disease, manifest by loss or disruption to the specialized brain cells in the body. One way in which we can treat such illnesses is by replacing lost cells with new ones generated in lab and transplanted into patients. However, a reliable source of those highly specialised cells is difficult to achieve. Therefore, such cell replacement therapies have not become used in the clinic. For multiple sclerosis, there is a need to produce oligodendrocytes – the cells that wrap an insulating like sheath around the nerves in the brain.

One approach to make the desired cells is by converting more abundant and easy to grow cells into the specialized cells. For example, clinicians could take some of patients' skin cells – which easily replenish and expand in the laboratory culture dish – program them into oligodendrocytes in the laboratory by resetting the patterns of gene activity: turning key specialised genes 'on' and the non-desired host genes 'off'. Some experimental studies have shown this is possible in the lab culture dish; however, the process is still largely inefficient, unreliable and costly.

In this work, we have investigated an emerging strategy to reset specialised cell types. This approach is based on taking control of the cells' genes by using artificially created transcription factors – the key proteins involved in switching on and off gene activity. Recently it became possible to produce such synthetic transcription factors (sTFs) using a new technology called CRISPR. Here we explore how well these technologies can work in neural stem cells and what the key parameters are for making them work effectively. We build a new set of tools to work with and deliver

these sTFs, and provide a proof-of-principle by taking control of genes involved in making oligodendrocyte progenitor cells.

Acknowledgements

Firstly, I would like to thank my supervisor Prof Steve Pollard. His insightful comments and guidance throughout the project was essential for its' success. Besides his sharp mind and extensive knowledge in the field, he also was understanding and fun to be around which made this whole journey much more enjoyable.

Secondly, I would like to thank everyone in our lab. Sabine and Harry were very important role models in my first year and helped me find my way around the lab as well as inspired me to work hard. Claudia, Ester, Pooran, Kirsty, Colin, Ute and Maria Angeles were amazing friends and I will always remember our time together. Raul, Vivian, Maria, Katarina, Carla and Steffi were great colleagues and I thank you for your help throughout the years.

Last but not least, I would like to thank my family who always were the source of motivation for me. They gave me strength and grit to pursue this and other endeavours. Most importantly, I would like to dedicate this work to my sisters Laura and Zivile who are and always will be my guiding stars.

List of abbreviations

AmpR	Ampicilin resistnace
BFP	Blue fluorescence protein
BMP	Bone morphogenic protein
BRE	TFIIB recognition element
BSA	Bovine serum albumin
BSD	Blasticidin
Cas9	CRISPR associated protein 9
cDNA	Complementary DNA
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR targeting RNA
dCas9	Dead Cas9
DCE	Downstream core element
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOX	Doxycycline
DPE	Downstream promoter element
DSB	Double-strand break
EGF	Epidermal growth factor
ESC	Embryonic stem cells
FACS	Fluorescence activated cell sorting
FGF	Fibroblast growth factor
FISH	Fluorescence in situ hybridization
GFP	Green fluorescent protein
gRNA	Guide RNA
HEK	Human embryonic kidney cells
HR	Homologous recombination
ICC	Immunocytochemistry
Inr	Initiator
iPSC	Induced pluripotent stem cells
IRES	Internal Ribosomal Entry Site
KRAB	Krüppel associated box
MEF	Mouse embryonic fibroblasts
MET	Mesenchymal-to-epithelial transition
mRNA	Messenger RNA
MTE	Motif ten element
NHEJ	Non-homologous end joining
NSC	Neural stem cell
OL	Oligodendrocytes

OPC	Oligodendrocyte precursor cells
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PDGFR α	Platelet-derived growth factor alpha-receptor
PEI	Polyethylenimine
PIC	Preinitiation complex
qPCR	Quantitative PCR
RFP	Red fluorescence protein
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase PCR
SDS PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
SHH	Sonic Hedgehog
sTF	Synthetic transcription factor
TAD	Transcription activating domains
TAF	TBP-associated factor
TALEN	Transcription activator-like effector nuclease
TBP	TATA-binding proteins
TCT	polypyrimidine initiator
tracrRNA	Trans-activating crRNA
TRE-3G	Tetracycline Response Element – Third generation
tRNA	Transfer RNA
TSS	Transcription start site
TSS	Transcription start site
VP16	Herpes simplex virus protein vmw65
VPR	VP64-p65-Rto
ZFN	Zinc finger nuclease

List of figures

Figure 1-1 Schematic of cell state changes during development or through experimental programming, focussing on the neural lineage.	25
Figure 1-2 Epigenetic modifications and nucleosome structure.	28
Figure 1-3 Two most common approaches to induce cell type change.	34
Figure 1-4 Graphical representation of CRISPR-Cas9 system and principles of DNA damage repair which are exploited for gene editing.	39
Figure 1-5 Schematic of a dCas9-VP160 sTF with its' partners and potential advantages of sTFs.	44
Figure 1-6 Structure of core promoters and graphical representation of transcription machinery assembly.	52
Figure 1-7 Generic summary of key markers and regulatory proteins during OPC specification and subsequent oligodendrocyte differentiation.	60
Figure 1-8 Graphical representation of main hypothesis.	65
Figure 3-1 gRNA placement for <i>Sox10</i> , <i>Olig2</i> and <i>Nkx6-2</i> gene activation.	81
Figure 3-2 gRNA-expression plasmid construction.	82
Figure 3-3 <i>Sox10</i> , <i>Olig2</i> and <i>Nkx6-2</i> activation using sTFs in mouse neural stem cells (mNSCs; blue) and mouse embryonic fibroblasts (MEFs; red).	84
Figure 3-4 Comparing small and large pools of gRNAs in their ability to activate <i>Olig2</i> and <i>Nkx6-2</i> in mouse NSCs and MEFs.	87
Figure 3-5 Graphical representation of the image cytometry workflow.	90
Figure 3-6 <i>OLIG2</i> activation in HEK293T using multiple gRNAs.	93
Figure 3-7 HA-dCas9-VP160 levels in transfected HEK293T.	94

Figure 3-8 Workflow and optimization of RNA-FISH using <i>Gapdh</i> probes in MEFs.....	97
Figure 3-9 Co-RNA-FISH for <i>Sox10</i> and dCas9 in MEFs after activation of <i>Sox10</i> with sTFs.....	98
Figure 3-10 Introduction to PDGFR α -GFP (PG1.1) cell line.....	100
Figure 3-11 <i>Olig2</i> activation during PG1.1 differentiation to OPCs.	102
Figure 3-12 <i>Olig2</i> activation during PG1.1 differentiation to astrocytes...	104
Figure 3-13 Screening of new transcription activating domains (TADs)...	107
Figure 3-14 gRNA placement for Sox9 gene repression.	110
Figure 3-15 <i>Sox9</i> repression in SOX9-V5 cell line.....	111
Figure 4-1 Lack of effective co-transfection poses a hurdle for multiplex delivery of sTFs.....	117
Figure 4-2 Summary of an Extensible Mammalian Modular Assembly (EMMA) system and the process of assembly.	118
Figure 4-3 EMMA plasmid assembly using fluorescent proteins and localization tags.	121
Figure 4-4 Making and validating empty All-in one (Ai1) vector for multiplex gene activation.	122
Figure 4-5 Adaptation of EMMA system for assembly of plasmids containing multiple gRNAs and dCas9-VP160.....	125
Figure 4-6 U6-tRNA-gRNA-tRNA array system for multiplex gene regulation.....	128
Figure 4-7 U6-gRNA array system for multiplex gene regulation.	130
Figure 4-8 Sequencing of 2SON clones after re-transformation.....	131

Figure 4-9 System for simultaneous gene activation and repression.....	133
Figure 4-10 Making VRER dCas9 mutants.....	135
Figure 4-11 Validating VRER exclusivity using dCas9-VP160.....	137
Figure 4-12 <i>Sox9</i> repression using VRER-dCas9-KRAB in SOX9-V5 NSC line.	139
Figure 4-13 <i>Olig2</i> repression using VRER-dCas9-KRAB in PG1.1 cell line.	140
Figure 4-14 Evaluating SOX10 antibodies.....	142
Figure 4-15 Strategy for V5 tagging of C-terminus of <i>Sox10</i> using CRISPR- Cas9 protein and in vitro transcribed gRNAs in PG1.1 NSC line.....	144
Figure 4-16 Derivation of PG1.1 SOX10-V5 clonal cell lines.	145
Figure 4-17 Derivation of BL6 SOX10-V5 clonal cell lines.	147
Figure 5-1 PDGFR α -GFP (PG1.1) reporter cell line.....	151
Figure 5-2 Construction of PG1.1 cell line with DOX inducible <i>SOX10</i> and <i>OLIG2</i> (PG1.1-SO).....	153
Figure 5-3 Construction and characterization of PG1.1-SO clonal cell lines.	155
Figure 5-4 Construction of stable PG1.1 cell lines containing 3SON sTFs (PG1.1-3SON).	157
Figure 5-5 Effect of <i>SOX10</i> and <i>OLIG2</i> overexpression and <i>Sox10</i> activation on mouse neural stem cells in self-renewal conditions.	160
Figure 5-6 Effect of <i>SOX10</i> and <i>OLIG2</i> cDNA overexpression and <i>Sox10</i> activation on mouse neural stem cells in OPC differentiating conditions.	162

Figure 5-7 Differentiation of PG1.1 and PG1.1 SO c4 NSC to OL lineage.	164
Figure 5-8 Characterizing BL6 cell line differentiation potential in different differentiation inducing growth mediums.....	166
Figure 5-9 Constructing stable BL6 cell lines expressing sTFs targeting <i>Sox10</i> (BL6-S10).....	167
Figure 5-10 Effect of <i>Sox10</i> activation on NSC differentiation to OL lineage.	169
Figure 5-11 Generating MEF lines from mouse embryos.	171
Figure 5-12 MEF transdifferentiation to OPCs/OLs after <i>Sox10</i> , <i>Olig2</i> and <i>Nkx6-2</i> activation with 3SON.....	174
Figure 5-13 MEF transdifferentiation to OL after <i>Sox10</i> , <i>Olig2</i> and <i>Nkx6-2</i> activation with 3SON.....	175
Figure 5-14 Optimized MEF transdifferentiation to OL after <i>Sox10</i> , <i>Olig2</i> and <i>Nkx6-2</i> activation with sTFs.	176
Figure 6-1 Working hypothesis for how multiple gRNAs enhance transcription activation.....	185

List of tables

Table 2-1 | Summary of primary antibodies used.71

Table 2-2 | List of TaqMan assays used.73

Table 2-3 | List of vectors used in study.....77

Table 8-1 | List of gRNAs used in this study and their sequences.....235

Table 8-2 | List of custom-made RNA-FISH probes and their sequences. ..237

Table 8-3 | List of PCR primers used for construction of Ai1 vectors238

Table 8-4 | List of PCR primers used for site-specific mutagenesis.....239

1 Introduction

1.1 Overview

Synthetic biology is providing a wealth of new tools to engineer mammalian cells. Some of the key areas under investigation are genetically encoded computing circuits, programming of cellular type and functions, and fully synthetic artificial chromosomes that can carry large cargoes (Li & Weiss 2017). These advanced technologies will underpin a myriad of new applications across diverse fields, e.g. food production, fuel production and medicine (Khalil & Collins 2010).

Efficient methods to rapidly and robustly experimentally direct mammalian cell type identity is a key challenge of stem cell biology. This is vital if new cell therapies are to become a common practise in the clinic, as current approaches are inefficient and unreliable. Given the increasing burden of neurodegenerative disease across the globe, neuroscience is one area that would particularly benefit by such developments. Routine production of neuronal or glial cell types in the laboratory is a major aspiration of the field. Two examples are dopaminergic neurons to treat Parkinson's disease and oligodendrocyte progenitors to treat demyelinating diseases such as multiple sclerosis (Ben-Hur 2011). However, current strategies to generate such cells through human ES or iPS cell differentiation involve long, complex and unreliable de-differentiation and differentiation protocols. These methods normally try to recapitulate developmental events *in vitro*.

What determines the identity of one cell type versus another? At the fundamental level a cell type is defined by its patterns of gene expression; it is a long

checklist of which genes are ‘ON/OFF’. So, what if we developed experimental approaches that enable simultaneously to switch genes ON and OFF? Could this be used to shift cell type identity? Based on the findings from the field of nuclear reprogramming – either by nuclear transfer, cell fusion or overexpression of master regulatory transcription factors – this would seem possible (Sparman et al. 2009). However, these reprogramming strategies are highly inefficient. Recent developments in synthetic transcription factor technologies suggest this might provide an alternative approach. This is the major topic of this thesis.

CRISPR technologies are now widely used to edit mammalian genomes; i.e to cut, tag, repair and insert genetic material. However, they can also be repurposed as new synthetic transcription factors. Nuclease dead variants of Cas9 (dCas9) can be fused to a transcriptional regulatory domain. These have been shown to be able to activate and repress target genes in mammalian cells. Although this technology is still in its infancy, it is likely to become an important tool in mammalian stem cell and synthetic biology. There could be important applications for generating cells for cell replacement therapies.

Our goal is to test the effectiveness of these synthetic transcription factors (sTFs) in programming and reprogramming neural cell types. In this introduction, I will discuss relevant areas; first, existing methods to program cell types; second, genome editing technologies, with an emphasis on the CRISPR system and its repurposing to make synthetic transcription factors; and finally, the molecular and cellular mechanisms that underlie neural stem cell self-renewal and differentiation to oligodendrocytes.

1.2 Programing and reprograming cell types

Efficiently controlling mammalian cell type identity has been a major objective of stem cell biologists and significant progress has been made (J. Xu et al. 2015). In this thesis, we will be using following terminology, which is mainly based on changes that occur in respect to cell potency. These are: programing (or differentiation) and forward programing, in which cells transit from higher potency to lower potency through more or less developmentally-relevant stages; reprograming, which is conceptually similar but mechanistically different from de-differentiation, where cells moves from lower potency to higher potency; direct reprograming (also known as transdifferentiation or direct lineage conversion) in which one somatic cell (low potency) can be directly converted to another somatic or progenitor cell (usually of different lineage but not necessarily) without going back to pluripotency (Fig.1-1). However, these terms are often used loosely and mean different things to different people. When applicable, we will present those in relevant parts of this introduction and try to be specific in terms of this project. This thesis is focussed on differentiating neural stem cells to oligodendrocytes (programing) and transdifferentiating fibroblasts to oligodendrocytes (direct reprograming).

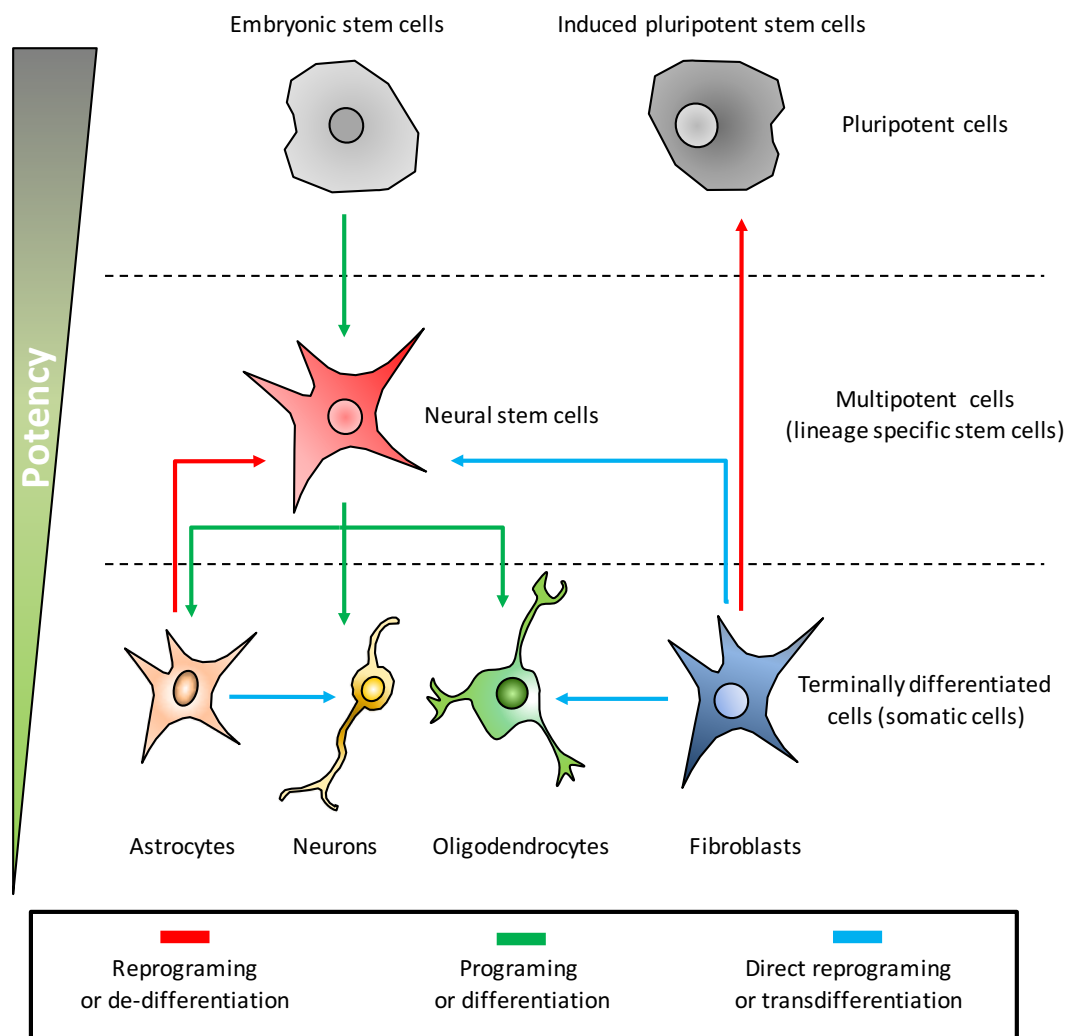


Figure 1-1 | Schematic of cell state changes during development or through experimental programming, focussing on the neural lineage.

The green triangle indicates the diminishing potency of cells during development.

1.2.1 Cell differentiation during development and *in vitro*

Cellular differentiation is defined as the process during which cell becomes more specialized to perform specific functions. Pluripotent cells of the blastocyst (equivalent of embryonic stem cell cultures) transit into multipotent tissue-specific progenitors (e.g. neural stem cells) and these generate terminally differentiated progeny (e.g. oligodendrocytes). As differentiation is accompanied by a loss of potency to generate different cell types, it is important that a subset of adult cells is retained in a stem cell state. These adult tissue stem cells continue to proliferate and differentiate in order to maintain tissue homeostasis or provide a source of new and healthy cells to replace damaged or diseased tissues. It was thought that this process is irreversible; essentially unidirectional. While this is true under most circumstances, it became evident from decades of research into nuclear reprogramming that this can be reversed experimentally – most clearly demonstrated by the studies of Shinya Yamanaka for the production of iPS cells (Takahashi et al. 2006).

Differentiation is often guided by external inductive signals, such as growth factors, hormones and morphogens. Cells exposed to such molecules activate various signalling pathways which then can result in cell type change; often these are signals that are used during normal development and to some extent are recapitulated in the culture dish. The timing, levels, duration and combination of many inductive cues determines the precise trajectory of differentiation. Detailed differentiation protocols have been developed for mouse embryonic stem cells, enabling differentiation to a range of specific lineages, including derivatives of all three major germ layers: reviewed in (Murry & Keller 2008); iPSCs (Mertens et al. 2016) and NSCs (Massirer et al. 2011; Glaser et al. 2007). Lastly, differentiation process can be facilitated by

overexpression of specific master regulators, a process termed forward programming. Such transitions have been demonstrated to be faster and give more homogenous outcome compared to simple addition of differentiation inducing culture medium (Pawlowski et al. 2017). It is not clear if forward programming is just an enhancement of differentiation process or if it relies on novel intermediate stages and transcriptional networks.

In addition to external stimulus, another important consideration is internal capacity of cell to respond appropriately – their competence. Such cell autonomous programs unfold during development with defined timing and this influences the way in which signalling pathways are read out (Singh et al. 2015; Llorens-Bobadilla et al. 2015). The ability of a cell to read out signals is heavily influenced by its lineage – events from earlier in development are recorded as epigenetic memory limiting options for further shifts in cell type. Therefore, while regulation at transcription level (via transcription factors) is crucial for understanding cell competency and process of cell type conversion, epigenetic control and mechanisms behind it are as, if not more, important to consider. In fact, epigenetic information works hand-in-hand with DNA sequence (genetic information) to regulate gene transcription and therefore cell identity.

Epigenetic control can be exerted both via modifications to DNA and histones - proteins wrapped around DNA (Fig. 1-2). While DNA directly can only be modified by methylation, which is often associated with transcription repression, histone tails and core proteins have a larger number of possible modifications. Histone acetylation, phosphorylation and ubiquitination has been linked to transcription activation, while sumoylation is linked to repression. Importantly, histone methylation has been linked

to both transcription repression and activation depending on the context of present enzymes and other modifications (Zhang & Reinberg 2001). All these modifications, essentially create a local gene environment that determines what kind of factors can access DNA at specific locations and time. In such a way, epigenetics plays a key role in cell's competence and identity, as it influences the potential for genes to be activated or repressed.

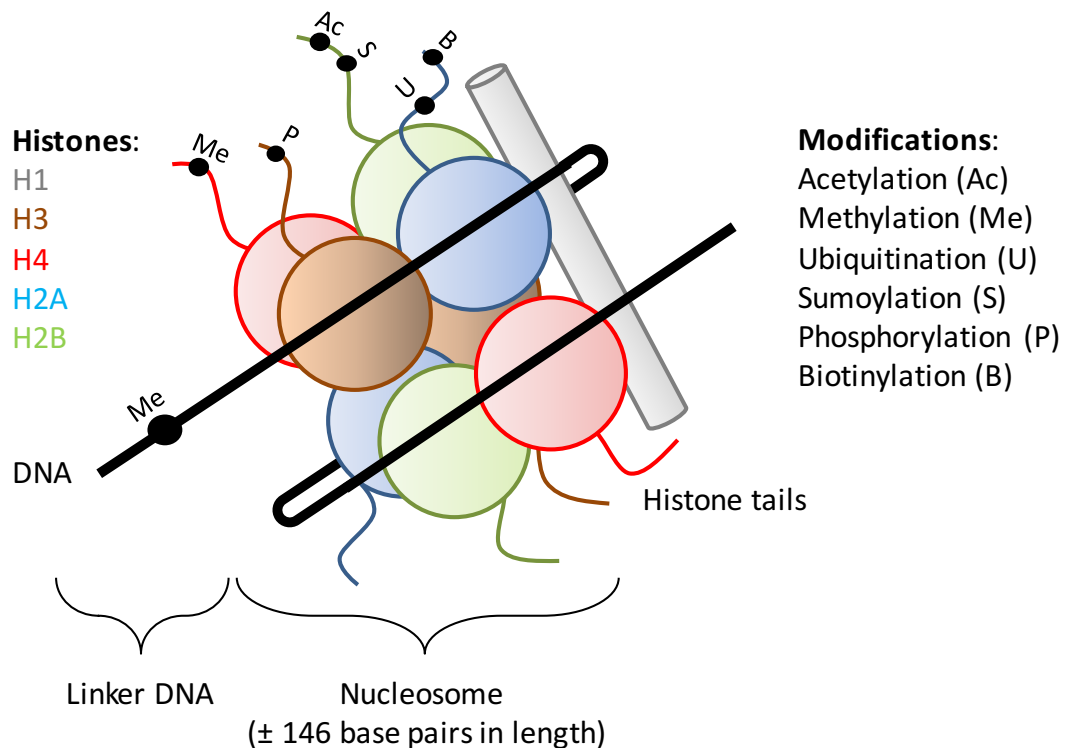


Figure 1-2 | Epigenetic modifications and nucleosome structure.

Nucleosome, a fundamental DNA organization unit, is composed of eight core histone proteins, their tails and additional linker histone subunit (H1). Such core is wrapped around (twice) by around 146 base pairs of DNA. Multiple epigenetic modifications can take place on histone tails as well as core proteins. Such modifications allow another layer of control over transcription.

Control of stem cell differentiation in the laboratory is a major challenge, but is essential for their application in regenerative medicine. Thus, understanding the transcriptional, signalling and epigenetic processes that control normal development is important, as this can help direct efforts to differentiate stem cells in the culture dish (Ebina & Rossi 2015; Neman & de Vellis 2012). Although this approach has a lot of potential, it is hugely limited by the need for complex cocktails of signals, patient-specific pluripotent cells, stepwise protocols and the long time required (often months) to transit from immature cells to the desired terminally differentiated cell type (Wang et al. 2013).

1.2.2 Reprogramming reverses differentiation to generate pluripotent stem cells

In the mid 20th century it was still unclear whether the process of differentiation was irreversible and whether it involved the removal of certain genetic sequences. This was tested directly by Gurdon et al. in 1962. In these experiments, Gurdon removed nucleus from *Xenopus* eggs by UV enucleation and then transplanted nucleus from somatic cell into these eggs (Gurdon J.B., 1962). Surprisingly, this gave fertile male and female adult frogs. This demonstrated for the first time that genetic information is not lost during differentiation and that somatic nucleus can become pluripotent in appropriate environment.

This and other work led to the studies by Yamanaka, who demonstrated reprogramming of the somatic cell to become induced pluripotent stem cells (iPSCs) by overexpression of four transcription factors Oct4, Sox2, Klf4 and c-Myc (known together as OSKM factors). From the perspective of cell potency, such reprogramming

can be classified as an example of complete de-differentiation where somatic or multipotent cells of specific lineage are turned back to pluripotent state.

During recent years the mechanisms behind iPSC reprogramming has been actively pursued. In the current model, two distinct phases of reprogramming have been described: stochastic initiation phase and deterministic maturation-stabilization phase (Buganim et al. 2013). During initial stages, cells undergoing reprogramming show increased proliferation, epigenetic and metabolic changes and initiation of mesenchymal-to-epithelial transition (MET). This early stage seems to happen in unpredictable and stochastic fashion and therefore be very inefficient. Later stage (stabilization of new phenotype) involves activation of key pluripotency genes, changes in cytoskeleton, chromosome re-organization and silencing of OSKM transgenes. During this stage, OSKM factors are able to bind and activate most of pluripotency-related genes (Soufi et al. 2012).

Development of iPSCs reprogramming protocols to a large extent solved the problem of a lack of patient-specific pluripotent cells and is therefore a very important advance for regenerative medicine. From a cell therapy perspective, while differentiation protocols are long and not completely understood, the main limitation usually was lack of starting material – pluripotent cells which could be forced to differentiate. With the development of iPSC, this major hurdle was overcome. iPSCs have been reported to be a viable source for generating specific somatic cells for transplantations (Li et al. 2015; Polentes et al. 2012). Despite these remarkable discoveries, the efficiency and fidelity of iPSCs has remained extremely low, and long incubation times in culture are required (Robinton & Daley 2012). Moreover, teratoma

formation remains an important risk if differentiation process is incomplete or reversed after transplantation.

1.2.3 Transdifferentiation: direct reprogramming of somatic cells

A third way of cell type programming is transdifferentiation, also known as direct lineage conversion or direct reprogramming; this involves a transition from one somatic cell to another somatic or multipotent cell of same or different lineage without returning to pluripotent state. Transdifferentiation was first reported in late 20th century by Davis et al. when he successfully converted mouse fibroblasts to stable myoblasts by overexpression of MyoD cDNA (Davis et al. 1987). To date, a number of somatic-to-somatic transitions have been demonstrated by other groups (Kikuchi 2015; Vierbuchen et al. 2010; Vierbuchen & Wernig 2011). One good example of transdifferentiation has been reported by Yang et al. in 2013 study where they successfully converted rat fibroblasts into OPC-like cells. By overexpression of just three master transcription factors (*Sox10*, *Olig2* and *Zfp536*), they managed to obtain, on average, 15.6% O4-positive cells at the end of 20 days long protocol. These cells were capable of differentiating into CNP-positive cells and myelinating neurons when transplanted into mice incapable of endogenous myelination (Yang et al. 2013). Excellent review on the historical progress of transdifferentiation has been written by Xu and colleagues (Xu et al. 2015).

Transdifferentiation mechanisms are still poorly understood but some insights have been made. Firstly, transdifferentiation can either happen directly or involve initial step of de-differentiation. For example, during lens regeneration in adult newt, pigment epithelium cells initially show signs of de-differentiation as demonstrated by

the loss of pigment and increased proliferation. This is followed by activation of various genetic programs (similar to ones used in lens development) and cells' transdifferentiation (Tsonis et al. 2004). One difficulty in such cases is the similarity of the transdifferentiation process to that of de-differentiation followed by the differentiation. The clear distinction is hard to make. On the other hand, transdifferentiation from fibroblasts to muscle cells by overexpression of MYOD seems to avoid intermediate stem cell-like proliferative stage (Davis et al. 1987). Instead, transdifferentiating cells seem to enter into unnatural intermediate state. As work in B cell transdifferentiation has shown, this stage partly reassembles original and partly final cell (Xie et al. 2004). Interestingly, it is still not clear how much different is this intermediate unnatural state to that intermediate phase observed during reprogramming. Both seems to possess a mixture of original (somatic) and new (pluripotent in case of reprogramming, and somatic in case of transdifferentiation) genes. Mechanically, it could be that reprogramming and direct reprogramming work in a similar fashion. Secondly, epigenetic remodelling is an important process for transdifferentiation. It has been shown that transdifferentiation requires pioneering transcription factors to directly access nucleosomes and then coordinate recruitment and binding of secondary factors (Iwafuchi-Doi et al. 2014). Thirdly, it seems that transdifferentiation might be two-step process: initiation followed by maintenance of a new cell fate. Some of the evidence suggesting this comes from hepatocyte transdifferentiation to biliary epithelial cells. While cells which completed transdifferentiation process stay as epithelial cells, other cells which are still undergoing the process can return back to hepatocyte fate if stimulus is removed (Tarlow et al. 2014). Lastly, while cell division seems to be a critical requirement and

an enhancing condition during iPSCs reprogramming (Hanna et al. 2009), it is not required during transdifferentiation of fibroblasts into neural cells (Iwafuchi-Doi et al. 2014; Fishman et al. 2015).

While transdifferentiation might become very important method for some applications, it comes with its downsides. For example, while we can easily expand pluripotent cells (and therefore generate enough of them to be used for cell therapy), expansion of somatic cells is more difficult. Such limitation would mean that potentially reprogramming (dedifferentiation) followed by differentiation might be better approach for off-the shelf cell therapies. That being said, transdifferentiation has a lot of potential for *in vivo* cell modulation and reprogramming. For example, *in vivo* reprogramming of astrocytes (or other supportive and replicative cells) to OPCs in MS patients could potentially treat patients without a need to take cells out for *in vitro* culture for processing.

1.2.4 Master transcription factors and their mode of action.

Common to each discussed transition is the clear need to trigger new patterns of gene expression and associated gene regulatory networks. These are controlled by cell type specific transcription factors – often termed master regulators. Be it by activation of these master regulators by growth factors, or direct overexpression of them (Fig. 1-3), these transcription factors are central to defining cell identity in every cell type transition. Therefore, it is important to understand how these transcription factors operate and what makes them unique.

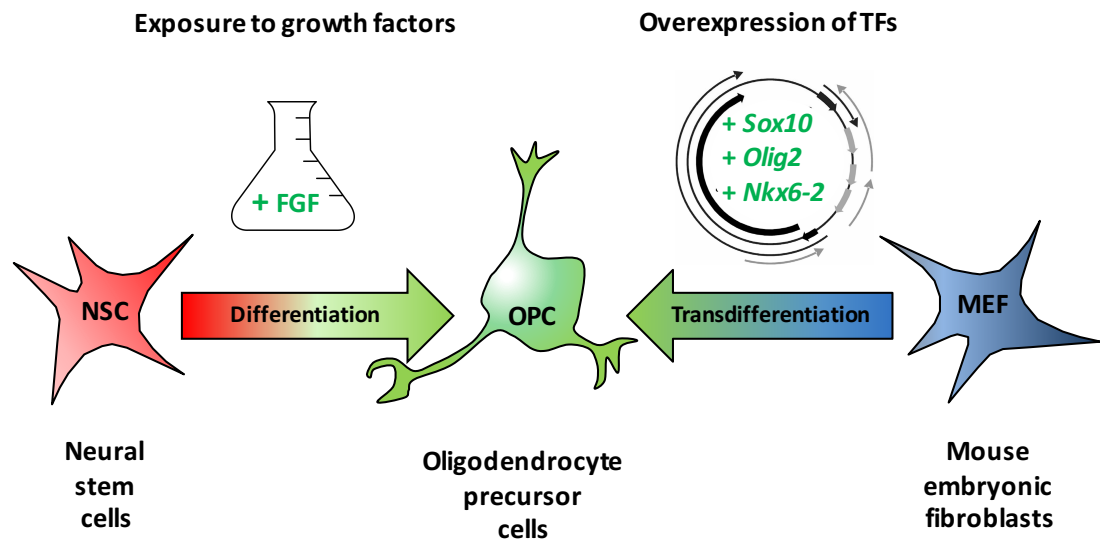


Figure 1-3 | Two most common approaches to induce cell type change.

Most commonly, programming and reprogramming events are achieved either by using a cocktail of specific growth factors (which when activate specific master regulators) or by overexpressing key master regulators directly.

All of the discussed transitions have been successfully induced by the overexpression of some key master regulatory genes; reprogramming (Takahashi et al. 2006); transdifferentiation (Najm et al. 2013; Yang et al. 2013); differentiation (Yao et al. 2014; Maire, Buchet, Kerninon, Deboux, Baron-Van Evercooren, Nait-Oumesmar, et al. 2009; Liu et al. 2007; Wang, Suyog U. Pol, et al. 2014). Although the mechanisms are probably somewhat unique for each conversion but some light has been shed on the general principles of how these master regulators generally work.

Firstly, transcription factors that will be able to drive the transition are usually specific to the cell lineages involved. Most studies investigating new cell type changes usually begin by screening panels of lineage-specific master regulators to identify the optimal combinations. Clear examples are MYOD in muscle transdifferentiation, SOX10 in oligodendrocyte transdifferentiation, OCT4 in iPSC reprogramming. Secondly, many of master transcription factors involved in cell type changes act as pioneer factors and are able to bind some of closed chromatin regions (Iwafuchi-Doi & Zaret 2016). These factors can bind partially repressed low-signal chromatin by binding some of the surface motifs on nucleosomes, displacing some of the linker histones and making nucleosomes accessible in such a way (van Oevelen et al. 2015; Soufi et al. 2015). One excellent study have investigated how Ascl1 acts as pioneering factor in first 48 hours post its expression (Wapinski et al. 2017). This group reports, based on ATAC-seq data, that nucleosomes are shifted from Ascl1 target sites in first 36 hours. Importantly, all of the nucleosome forms (mono-, di- and tri-) are affected. However, even pioneer factors do not seem to be able to target actively repressed chromatin regions (Iwafuchi-Doi & Zaret 2016). Thirdly, each master regulator might play a distinct role during different stages of programing. For example, during

transdifferentiation from MEFs to neurons – triggered by overexpression of *Ascl1*, *Brn2* and *Myt1l* – *Ascl1* plays a key pioneering role and is able to induce immature neurons by itself. Meanwhile, it has been shown that *Brn2* only has a supportive role and contributes to the neuronal maturation stage (Vierbuchen et al. 2010).

Another example comes from OSKM factors. *c-Myc* has been shown to act as enhancing factor which is replaceable (Nakagawa et al. 2007) or even redundant (Wernig et al. 2008) while *Oct4* seems to be critical: it cannot be removed or substituted (Feng et al. 2009). Consequently, it becomes clear why only having all required factors together allows an efficient cell type transition. Fourthly, both timing and levels of master transcription factors are critical. In the study by Sokolik et al., it has been showed that unless overexpression of *Brn2* is 100-fold or more over the basal line, mESCs do not differentiate to NSC (Sokolik et al. 2015). Moreover, if >100-fold overexpression is not maintained for more than 5 hours, *Nanog* does not get repressed and differentiation cannot occur. The authors hypothesize that such restrictions allow cells to distinguish signals to differentiate from transcriptional noise. Other groups have reported that iPSC reprogramming efficiency is highest if OCT3/4 levels were high and SOX2 levels were low (Papapetrou et al. 2009). All of these are important lessons which needs to be considered when optimizing new ways for inducing cell type conversions.

1.2.5 New ways of inducing cell type changes.

New experimental approaches to steer cell fate specification are also being developed, such as use of small pharmacological agents (Shi et al. 2008; Hou et al. 2013) and miRNAs (Zhao et al. 2012; Kuypers et al. 2016; Zhao et al. 2010; Budde et

al. 2010). Lately, a new method which makes use of synthetic transcription factors (sTFs) has been successfully tested (Black et al. 2015). This offers the exciting prospect of more fine-tuned control of target gene expression patterns. This is a major focus of thesis and is discussed in detail in the next section.

1.3 Genome editing technologies and synthetic transcription factors

1.3.1 Introduction to genome editing tools – Zinc finger nucleases, TALENs and CRISPR-Cas9

Genome editing is one of the most exciting new technologies now used widely by biologists to systematically dissect the functions and interactions of various genes and proteins. It enables precise engineering of DNA sequences at any desired location in the genome. The emergence of effective genome editing required developments in designer nucleases – also known as synthetic nucleases. These can be ‘engineered’ to bind at specific sites in the genome and introduce double-strand breaks (DSBs). DSBs are typically repaired by the non-homologous end joining (NHEJ) pathway, which is error-prone, and therefore can introduce random mutations at the target site. Alternatively, homologous recombination (HR) can be used in which a DNA template replaces the region surrounding the DSB (Fig. 1-4C). There are a number of different formats, or architectures, upon which such synthetic nucleases can be built. Zinc finger nucleases (ZFN) were the first designer nucleases to emerge. A key technological development was the discovery by Srinivasan Chandrasegaran that fusion of the DNA binding domain of a ZFN to the FokI restriction enzyme can create a novel fusion protein with specific DNA binding and effective DNA endonuclease activity (Kim et al. 1996). These new synthetic sequence-specific nucleases bound DNA via zinc finger

domains, in which each domain recognized 3 base pair spans; as the ‘rules’ for predicting target sequence of synthetic domains were developed, it became possible to target this to specific sites. However, ZFN exhibited high off-target effect and cytotoxicity (Kim et al. 2009) and the design and construction of the DNA binding domain was often difficult.

Transcription-activator-like effector (TALE) proteins emerged as an alternative architecture. TALEs when fused to the FokI domain can create nucleases named TALENs (Boch et al. 2009). TALENs have advantage over ZFNs because each TALE domain recognizes individual base pairs and therefore have higher target-specificity and lower off-target effects and are easier to design. Furthermore, while early TALEs were unable to bind methylated DNA, development in the field allow these nucleases to target even methylated DNA (Scott et al 2014). Unfortunately, TALENs are still demanding to construct at scale, as each time new nuclease is constructed a modular assembly of repeated domains is required. So, protein engineering and the need to create unique plasmids was a hurdle. Although a number of protocols have been designed to make TALE assembly more efficient and straightforward (Cermak et al. 2011; Reyon et al. 2012), this still requires a significant investment and expertise.

Most recently another genome editing platform, termed CRISPR/Cas9, has rapidly overtaken ZFN and TALENs as the preferred technology for genome editing (Lander 2016). Cluster of regularly interspaced palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) (herein referred to as CRISPR) is responsible for bacterial immune response to viral infection (Mojica et al. 2005; Bolotin et al. 2005; Gasiunas et al. 2012; Jinek et al. 2012). It was fully repurposed for use in mammalian

cells for genome editing in 2013 (Cong et al. 2013; Mali, Esvelt, et al. 2013). In the following sub-section I discuss CRISPR technologies in detail as they form the core technology used in this thesis.

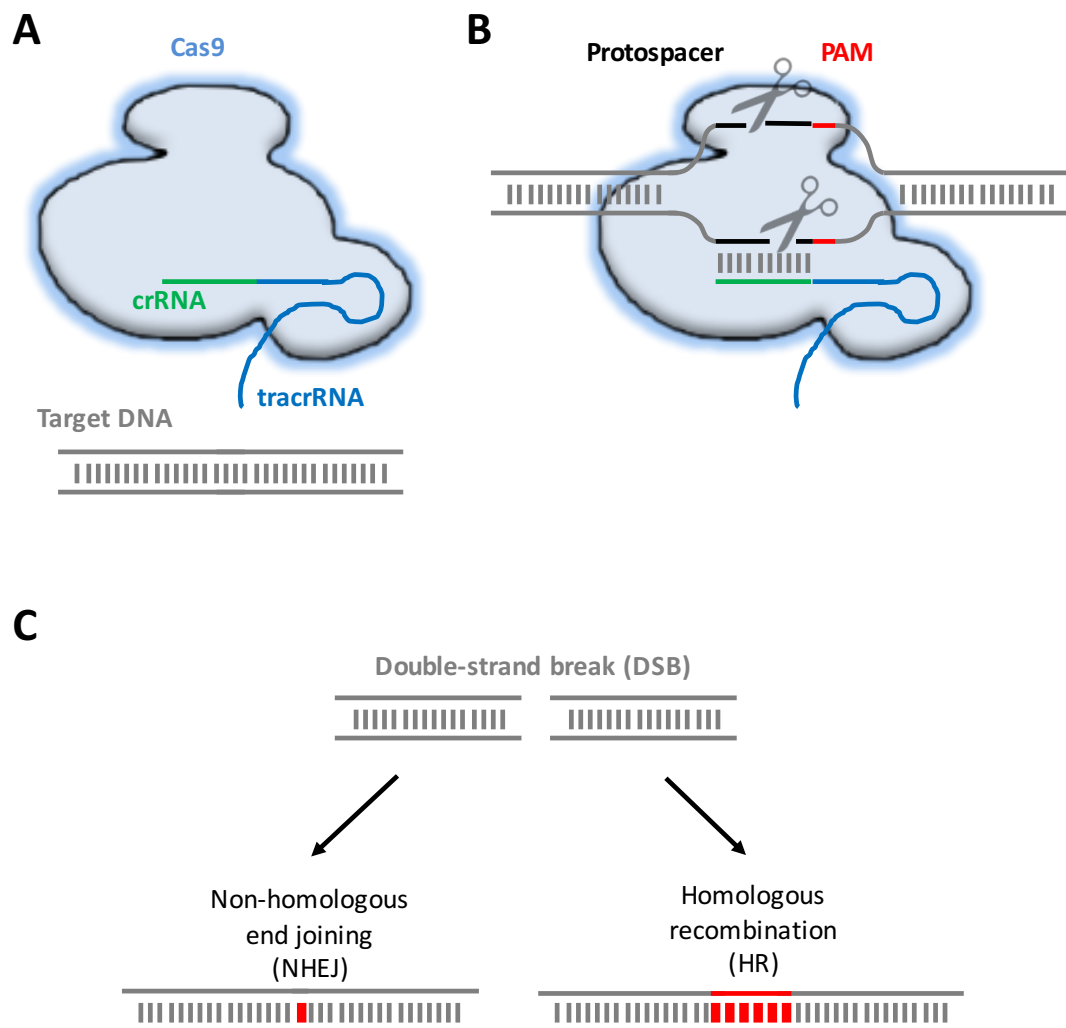


Figure 1-4 | Graphical representation of CRISPR-Cas9 system and principles of DNA damage repair which are exploited for gene editing.

(A) Cas9 nuclease forms a complex with a guide RNA and scans the genome for complementary DNA sequence. (B) Upon detecting sequence of interest, Cas9/gRNA opens double helix and initiates DNA strand cutting using two separate nuclease domains. This generates a double-strand break at the site. (C) Graphical representation of how double-strand breaks can be repaired. NHEJ results in insertion/deletion mutations at the loci of interest; alternatively, HR repair allows more elaborate modification (i.e. tagging) of the site by replacement of sequence with exogenous DNA. New techniques have been developed to favour one or another repair mechanism, allowing greater control over genome editing efficiency.

1.3.2 Biology of CRISPR-Cas9

The mechanism behind CRISPR-Cas9 has been intensively studied for over a decade. The system is composed of a Cas9 RNA-guided nuclease and separate guide RNA (gRNA), which itself is a combination of crRNA (CRISPR targeting RNA) and tracrRNA (trans-activating crRNA). crRNA is complementary to a target sequence, or protospacer, and protospacer has to have an appropriate PAM (protospacer adjacent motif) sequence at its 3' end to be a valid target for CRISPR-Cas9 (Fig. 1-4).

Cas9 protein forms a complex together with gRNA (Tsai & Joung 2016). This is followed by the search for matching PAM sequences across the genome. Each time a PAM sequence is encountered, Cas9-gRNA loosely binds the DNA but does not initiate the cutting. Cas9 therefore quickly scans extremely large spans of DNA. If the sequence adjacent to the PAM matches the crRNA then Cas9 protein undergoes conformation change, which significantly increases binding affinity of Cas9-gRNA to the target sequence and leads to cleavage of the DNA. Importantly, the correct PAM sequence is used as a checkpoint at multiple stages of this process – from initial search to the process of cleaving DNA (Sternberg et al. 2014). One of the key steps for this process success is an ability of Cas9 to reach and bind even hard-to-access DNA sequences. Large parts of the genome are constantly in highly compact chromatin state, known as heterochromatin. Fortunately Cas9 can indeed access and cleave at such sites (Polstein et al. 2015). However, there is evidence that nucleosomes can interfere with Cas9 functionality (Hinz et al. 2015; Horlbeck et al. 2016; Isaac et al. 2016), which could partly explain why some gRNAs have such variable functionality in distinct cell types. For example, in the study by Horlbeck and colleagues it has been reported that Cas9 not only cannot cut DNA if nucleosomes are introduced, but dCas9 cannot even

bind such DNA (Horlbeck et al 2015). If nucleosome remodelling enzymes are added (such as γ Chd1), Cas9 recovers its ability to cut such DNA, demonstrating strong negative impact of nucleosomes on Cas9 functionality.

Compared to TALENs and ZFNs, CRISPR-Cas9 has some salient features that are experimentally desirable. CRISPR-Cas9 is a much more scalable tool as it does not require construction of new proteins to target different sequences in the genome; new sites can be targeted by simply acquiring a different gRNAs. Such plasmids can be easily made and also can be amenable to large scale library preparations for pooled screens. While both TALE and ZFN cleavage requires dual binding as heterodimers, wild type Cas9 operates as nuclease (Kim 2016); this means there are fewer restrictions on choice of target sites.

To date, CRISPR has been used for a large number of different applications. Generally, these can be segregated into two main areas: (i) genome editing by introducing or removing DNA sequences (random point mutations, tags/reporters, landing pads, modified genes etc.) and (ii) various non-cleaving manipulations performed by dead Cas9 or dCas9 (Wang et al. 2016). dCas9 is mutant version of Cas9 whose both catalytic sites have been mutated to completely remove nuclease activity. dCas9 has been successfully used to image genetic locus (Chen et al. 2013), analyse protein-genome interactions (Fujita & Fujii 2013), and activate or repress gene transcription by acting as synthetic transcription factors (sTFs).

Understandably, CRISPR use for genome editing came with initial concerns about off-target effects when used in mammalian cells. These concerns stemmed from the speculations that bacterial protein would misbehave in context of much larger

mammalian genome and that monomeric nature of Cas9 nuclease would result in more off-target DSB. Great efforts have been made to explore these potential limitations (Slaymaker et al. 2016; Kim et al. 2015) and to create tools to circumvent such problems. One example is the creation of artificial nickase Cas9 by mutating one of its catalytic nuclease domains, such that only one strand is cleaved. This has been used to improve target specificity, as two nicks on each strand that are in close proximity (<100bp) are less likely to occur on off-target site compared to a single DSB. This has been demonstrated in mouse embryos by Shen et al. After targeting *Ar* gene with wild-type Cas9, at least three off-target sites have been cleaved. However, if nickase Cas9 and two sgRNA were used, no detectable off-target effects were seen (Shen et al. 2014). Similar results have been reported by other groups as well (Kim et al. 2015). That being said, other studies have reported that up to 60 non-specific cuts (of frequency as high as 10%) can be identified for some of the targets/gRNAs (Tsai et al 2016). Furthermore, current sequencing limitations might mask some very infrequent, but still therapeutically very important, off-target cut sites. That being said, to date it seems that off-target effects are not a major problem for use of CRISPR and especially not for CRISPRa (Cho et al. 2014).

1.3.3 Synthetic transcription factors

A synthetic transcription factor, or sTF, is an engineered artificial protein containing a DNA binding domain and separate effector domain which can activate or repress transcription. When these proteins are targeted to promoters or enhancers using engineered DNA binding domains (TALE-TFs or ZF-TFs) or gRNAs (CRISPR-TFs), gene transcription can be regulated. These sTFs can be deployed in various applications, from alleviating diseases by manipulating gene expression to inducing

cell type changes (Himeda et al. 2016). For example, sTFs have been used to upregulate dystrophin to treat Duchenne muscular dystrophy and upregulate γ -globin, which is epigenetically silenced in sickle cell disease (Gersbach & Perez-Pinera 2014). Zinc-finger sTF which activated VEGF, increased wound healing better in comparison to overexpression of VEGF (Rebar et al. 2002). Examples like these make this new field very promising for both basic biology studies and various industrial and biomedical applications.

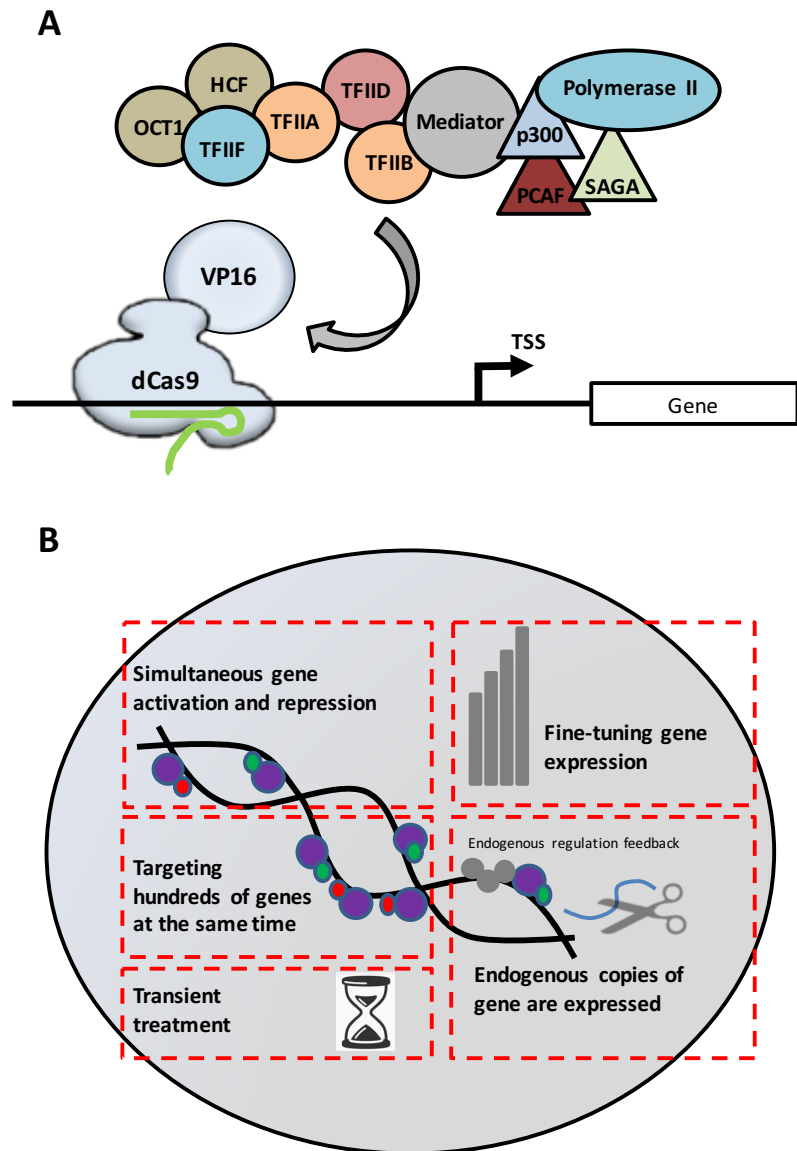


Figure 1-5 | Schematic of a dCas9-VP160 sTF with its' partners and potential advantages of sTFs.

(A) One of the most common dCas9-based sTFs used to date are dCas9-VP160. Its transcription activating domain VP16 is a viral protein that is capable of attracting various partners which are required for transcription initiation. (B) Summary of the potential advantages of using sTFs rather than cDNA overexpression in cell fate reprogramming.

ZFNs, TALEs and CRISPR have all been successfully used as sTFs (Sander & Joung 2014). There are a number of advantages of using sTFs to activate endogenous genes, rather than the typical approach of ectopic cDNA overexpression via plasmids or viral vectors (Fig.1-5B). Firstly, transcription activation with sTFs targets the endogenous gene, and therefore will be amenable to ‘natural’ levels and splice forms of the protein product. Secondly, it can be used to target several different genes for repression and activation in multiplex. Moreover, it should be possible to scale-up and modulate dozens and even hundreds of target genes at the same time. Thirdly, considering the extent of potential modifications (hundreds of genes repressed and activated), it might be possible that transient treatment would be sufficient to induce a permanent effect that is epigenetically stable. Fourthly, by making use of different gRNAs and effectors it would be possible to ‘tune’ activation levels to those desired if need be. Lastly, genes activated by sTFs would undergo endogenous feedback loops to have more physiological control over their activity, which is useful in some experimental or therapeutic contexts.

While the functionality of TALE-sTFs and ZFN-sTFs strongly depend on DNA-binding domain design, in the case of CRISPR-sTFs optimal design of gRNAs is critical, yet still poorly understood, aspect. Early work suggested that the best area to target sTF binding for achieving optimal activation is -450 to -50 base pairs upstream of transcription start site (TSS) and best region for repressing genes is -50 to +300 bp from the TSS, where 0 represents TSS (Gilbert et al. 2014). Nevertheless, efficient activation has been achieved targeting different areas and even enhancers (Hilton et al. 2015). Importantly, in some cases targeting intron sequences was essential for successful transcriptional activation, as in case of Myt1l in MEFs (Black

et al. 2016). These rules are therefore a good starting point, but are not universal or highly predictive. As it is evident, knowing and targeting genes around the TSS increases chances for successful transcription modification; however, defining the TSS is itself not a clear task. Not only do databases only give predication of TSS location but often different databases contradict each other. Moreover, a distinct TSS might be used in different cell types. Thus, context is important. It has been shown that FANTOM5 database gives the best predictions to date, and that optimizing TSS predication can give a 54% chance to identify a functional gRNA (Radziskeuskaya et al. 2016). That being said, other databases only lack behind by couple percents, suggesting that database choice is not main current limitation. Furthermore, tools have been developed to predict gRNA functionality based on its' sequence and this has shown to improve sTF function (H. Xu et al. 2015). Other studies have showed that gRNA sequence should not contain stretches of four (or more) uracils (Nielsen et al. 2013) and that repetitive nucleotides decrease gRNA functionality (Gilbert et al. 2014). Lastly, it has been proposed that pools of multiple gRNAs targeting the same proximal promoters trigger higher activation levels compared to use of individual gRNAs (Bikard et al. 2013; Mali, Yang, et al. 2013; Perez-Pinera et al. 2013). Interestingly, although significantly outnumbered, there are few studies providing evidence that using multiple gRNAs to target same gene has a negative rather than positive effect (Shao et al 2018). However, as we do not yet know if extremely high levels of activation are essential, or even desirable, to achieve functional changes (e.g. cell type programing), it remains to be seen if single or pooled gRNAs will be most effective. This is one of the central questions addressed in this thesis.

1.3.4 Effectors used with synthetic transcription factors

Another important, part of sTF design is the type of effector domain used. Neither TALEs nor CRISPR-Cas9 could act as sTFs if a potent transcriptional activator would not be attached. The strength of an effectors is potentially one of the most critical features of sTFs if they are to be used in programing.

A number of different transcription activating domains (TADs) have been successfully used to activate genes (Table 1-1). One of the most commonly used activation domains is the VP16 viral protein (Triezenberg et al. 1996) which interacts with multiple transcriptional components (Fig.1-5A), including TATA-binding protein, TFIIB and SAGA complex among many (Hall & Struhl 2002; Hirai et al. 2010). Although it seems to operate effectively in most instances, activation levels for some genes and in some cells are relatively low. Some improvement in activation potential was observed if multiple copies of VP16 were fused together, namely VP64 (4 copies of VP16) and VP160 (10 copies of VP16). To date, it seems that strongest activators are ones which combines a set of different effector, namely VPR and SAM (synergistic activation mediator) systems. A side-by-side comparison of these TADs has been conducted in the past (Chavez et al. 2016). While most of the work in the field is done focusing on these effectors, further effort is undertaken to investigate if activation potential can be increased even further by designing better effectors (Zhou et al. 2018).

dCas9-VPR, which stands for dCas9 tethered to VP64-p65-Rta fusion complex, was developed by Chavez et al (Chavez et al 2015). After screening more than 20 known domains involved in transcription, three most potent activators were

identified: VP64, p65 and Rta. When used together, this combination (VPR) strongly outperformed VP64 in activating a number of different genes in HEK293TT cells. Importantly, while *NEUROD1* activation in iPSC with dCas9-VP64 had no functional effect, activation with dCas9-VPR resulted in iPSC differentiation to neurons. This was one of the earliest examples demonstrating that the strength of the TAD is critical aspect to consider in functional experiments.

Another important development in TAD construction was assembly of SAM system by Konermann and colleagues (Konermann et al. 2014). This study reported two important observations. First, it has shown that TADs can be attached to gRNAs instead of dCas9 protein and that TAD placement have an impact on TADs' functionality. TAD attachment to gRNAs was made possible by the fact that a part of gRNA stem loop remains exposed even after dCas9/gRNA complex formation. By engineering RNA-protein interacting intermediates (MS2 protein), they have managed to attach TADs to gRNAs. Interestingly, this has shown slight improvement in gene activation potential as compared to same TAD being attached to dCas9 protein. Second, an important observation in this study was that although increasing the variation of TADs have a beneficial effect on gene activation potential, but only to the extent. Addition of p65 and HSF1 to VP64 increased the activation potential in comparison to VP64 alone, however any further modifications were not beneficial.

The repertoire of validated transcription repression domains (TRD) is not as diverse as activators. Currently, Krüppel associated box (KRAB) is the main repressive domain used (Gilbert et al. 2014). It has been showed to reshape the epigenetic landscape of the promoter and thereby repress transcription. Besides KRAB, the other common repressive dCas9 version is one without any domain. dCas9

can physically interfere with polymerase activity. Interestingly, use of CRISPR-dCas9 as a repression system seems to be more specific than RNAi-based repression, as single mismatch have detrimental effect for dCas9 but not for RNAi (Gilbert et al. 2014). This means that repression with dCas9 would have lower off-target effects compared to RNAi based methods for gene repression.

The capability to engineer functional sTFs strongly relies on the knowledge of biology of promoters and transcription machinery. This topic is discussed in the next section.

Effectors	Description	Reference
VP16	Herpes Simplex Viral protein which helps recruit host genes for viral gene activation	Triezenberg et al. 1996 Hall & Struhl 2002
VPR combination	VP64 (VP16 copies) p65 (part of NFkB TF) Rta (Epstein-Barr virus R transactivator; complements VP16) combination	Chavez et al. 2015 Hardwick et al. 1992
SAM combination	VP64, p65, HSF1 combination	Konermann et al. 2014
p300	Histone acetylation (H3K27)	Hu et al. 2014
KRAB	Induce heterochromatin formation by recruiting HP1, SETDB1 and other	Lupo et al. 2013 Schultz et al. 2002 Groner et al. 2010

Table 1-1 | Common effectors used in synthetic transcription factors.

1.3.5 Core promoters and transcription machinery in eukaryotes

The core promoter is the genomic location where DNA transcription initiates. Core promoters lie between -450 and -40bp from TSS, as deletion of this region has been shown to diminish promoter activity in reporter plasmids (Cooper et al. 2005). The size of the core promoter varies from around 40bp to 150bp and it can be composed of multiple elements. These include: TATA box, Initiator (Inr), TFIIB recognition element (BRE), polypyrimidine initiator (TCT), motif ten element (MTE), downstream promoter element (DPE), downstream core element (DCE). Each occupies a distinct position within core promoter and have different functionality (Fig.1-6A). Many of these elements are bound directly by TFIID, either via the TATA-binding protein (TBP) domains or TBP-associated factors (TAFs). Importantly, many promoters do not contain all of these elements. For example, TATA-boxes, which for long have been thought to be a key binding region in every promoter, are only present in approximately 30% of human promoters (Yang et al. 2007). While many functions that these elements perform remain unknown, our knowledge of them is increasing. For example, only Inr and TATA elements can initiate transcription by themselves (Smale et al. 1998), while DPE seems to act similarly to TATA in TATA-less promoters by locking PIC proximal to TSS.

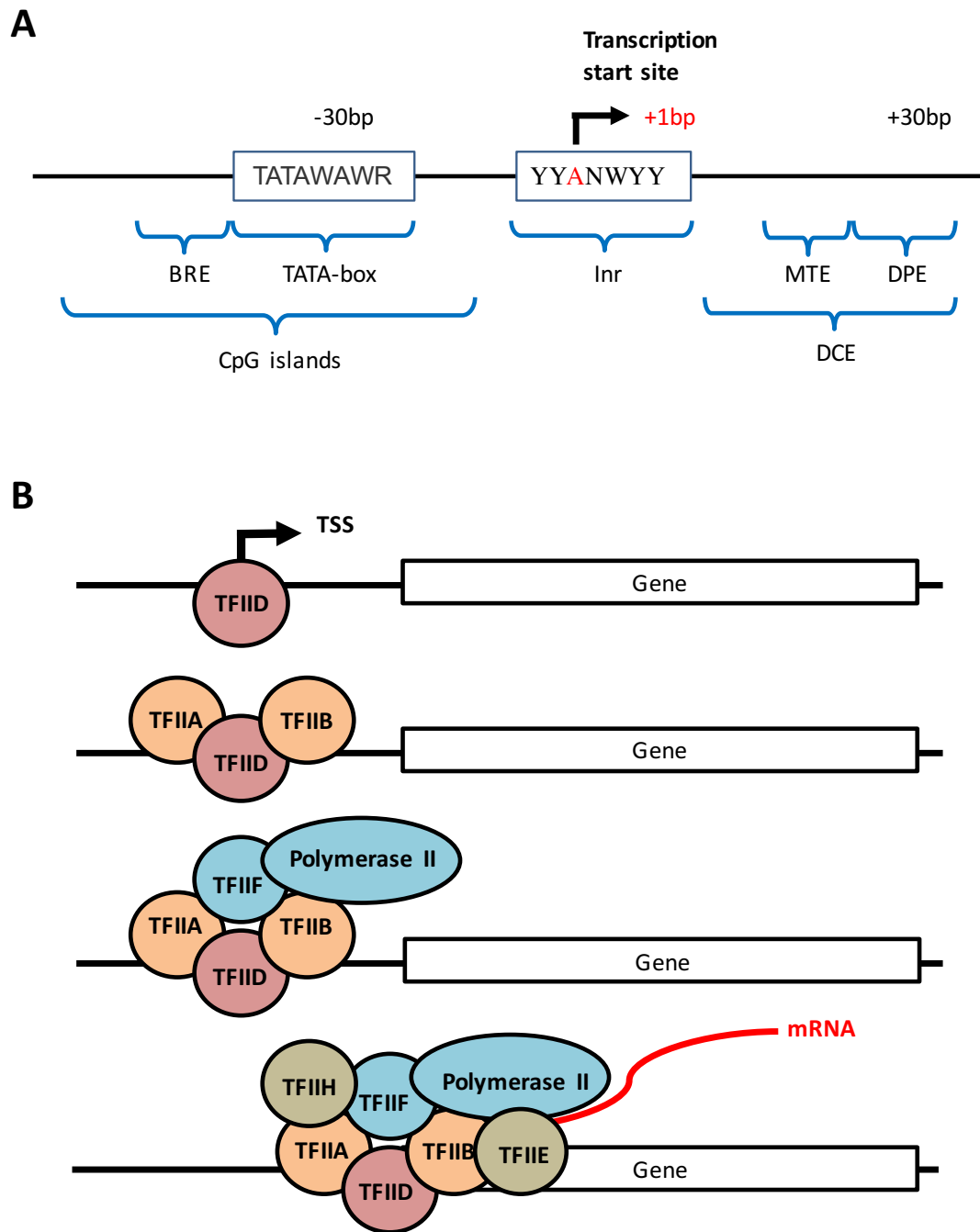


Figure 1-6 | Structure of core promoters and graphical representation of transcription machinery assembly.

(A) Core promoters consist of a number of elements, each playing a distinct role. However, there is variety in the composition of each promoter. (B) The transcription machinery assembles in stages. Partners bind in chronological order and various co-factors have to be present for transcription to start.

A second key feature of any promoter is transcription start site or TSS. This is a sequence within the promoter where transcription usually starts. If a promoter has an Inr element it usually lies within TSS. This region has been shown to be nucleosome-free (Nishida et al. 2006), which is important for transcription machinery and transcription regulation. Importantly, many promoters have more than one TSS and these can be distributed within 100bp region (Carninci et al. 2006; Suzuki et al. 2001). Often, promoters with multiple TSS's lack TATA-boxes, have CpG islands and are associated with constitutively expressed housekeeping genes; by contrast promoters with a unique TSS often have TATA-boxes and are associated with tissue-specific expressed genes.

It is clear that even core promoter architectures are very complex. Another layer of complexity and diversity is added when these promoters have to be tissue- or cell-specific. To achieve that, a large number of new factors and their binding sequences are introduced into promoter and enhancer sequences. Moreover, some genes are actually associated with more than one promoter (Sandelin et al. 2007). Such features allow production of different products (different introns/exon compositions), differential transcriptional response in different tissues etc.

In addition to promoters, another important element to consider for sTF functionality is the quality of the basal transcription machinery. Transcription in eukaryotes and prokaryotes has been widely studied. Eukaryotic cells have three different polymerases – Pol I, Pol II and Pol III – that are responsible for transcribing rRNA precursors, protein-coding regions and small non-coding RNAs, respectively (Sainsbury et al. 2015). Although transcription is controlled at many levels, the initiation stage seems to be the most critical and rate limiting. The general transcription

factor TFIID, which contains TATA-box binding protein (TBP) and TBP-associated factors (TAFs), assembles on promoter in the first step. Both TBP and TAFs are able to recognize many sequences present in core promoters and in such way TFIID is strongly involved in promoter recognition. Importantly, this step seems to be generic and TFIID has been shown to be involved in transcription of most genes in yeast (Huisinga & Pugh 2004). A second step involves TFIIA and TFIIB binding. This step seems to be optional in some cases, but it does increase affinity of TBP and stabilize the complex (Imbalzano et al. 1994). At this stage, PolII–TFIIF pre-formed complex combines with above proteins and DNA to form Preinitiation Complex (PIC). TFIIF plays a number of roles. It stabilizes many of the mentioned processes, prevents non-specific binding of DNA and also influences TSS selection (Ghazy et al. 2004). To complete “closed PIC”, TFIIH and TFIIIE binds as well. This is followed by use of ATP to open the DNA helix and create transcription bubble (aka “open PIC”). In presence of NTPs, transcription starts and many of these initiation factors are replaced by elongation factors (Fig.1-6B).

Although transcription is a precisely controlled process, it is often highly dynamic. This phenomenon is called transcription noise – a variability in gene expression between two genetically identical cells in the same population/conditions. There are a number of consequences of such behaviour. For example, noisy transcription can be a problem for cellular identity. Cellular identity is defined by a number of transcription factors and their continuous fluctuation would result in a lack of robustness in the cell type identity (Acar et al. 2005). Some of the explanations for transcriptional noise include transcriptional pausing/bursting, cell cycle effect (Zopf et al. 2013), stochastic nature of living systems (Raj et al. 2006) and slow transition from

inactivate to active promoters (Golding et al. 2005). It is most likely true that these and many other mechanisms play a role in transcriptional noise. How do cells solve this problem? It seems that specific rules have evolved on distinguishing signal from noise. Cells and their genetic networks responds only when context (other factors), levels and duration of signal is correct (Sokolik et al. 2015). These are important considerations for construction of effective sTFs that could be used to program patterns of transcription.

1.3.6 Synthetic transcription factors in cell type programing

sTFs have been used to achieve cell type programing either when used in combination with ectopic gene overexpression (Balboa et al. 2015) or when used alone (Black et al. 2015; Chakraborty et al. 2014). However, the process remains poorly understood. In a large majority of these studies only single gene has been activated to achieve reprograming and while this is sufficient for some cell type conversions, it will not be enough for other transitions. Efficiencies of reprograming are often not enhanced over cDNA overexpression. For example, a study by Black and colleagues demonstrated that sTFs can be used to convert MEFs into neurons by activation of *Myt1l*, *Ascl1* and *Brn2* (Black et al. 2016). Interestingly, the group has reported that both approaches (using cDNA overexpression or CRISPR-based activation of transcription) create a similar amount (2-4%) of *Tuj1* positive neurons. It remains unclear how widely such strategies can be used or if they are able to supersede cDNA overexpression as the main experimental approach for reprograming mammalian cell types

To make sTFs have broader utility a key missing capability is the use of multiplexing to test activation of many genes in parallel within the same cell. Multiplexing, in this context, means ability to target and activate multiple genes at the same time in the same cell; too often studies report activation of genes within the population, but it remains unclear if this means each sTF works well in individual cells. There are a number of reasons multiplexing is important. Firstly, some programming events can only happen if multiple genes are involved (like in iPSC reprogramming using OSKM). Secondly, if very large numbers of genes would be activated, efficiency of the programming might increase. After a cell is exposed to 3 or 4 reprogramming factors, they activate a number of downstream genes, which then activate other genes and so on until cells' genetic profile starts to resemble that of different phenotype. In theory, it might be possible to activate all of these genes simultaneously and in such way, force a more rapid lineage conversion.

A further reason why sTFs might have lot of potential in forcing cell type change is because repression of endogenous genes can be triggered at the same time as activation of other gene. There is an expanding list of evidence demonstrating that repression of genes can be as potent method for cell type programming as activation/overexpression is (Ikawa et al. 2015; C. D. Pozniak et al. 2010a). Importantly, repressors might be even more important than activators in some cases. One of the instances when this might be a case is when current cell type is maintained by strong (potentially pioneering) network of transcription factors. For example, *Ascl1* is a transcription factor involved with neuronal lineage (Shi et al. 2016). *Ascl1* is also able to bind closed chromatin (Raposo et al. 2015). As a result, it would be hard to repress any genes downstream of *Ascl1*, as *Ascl1* would be able to bind and re-activate

these genes. However, the pioneer activity and cell-type maintenance capability of Ascl1 would not matter if it would be directly repressed itself. Therefore, one can argue that some programming events might happen more effectively in the context where strongest cell type maintainers are repressed. Another example when repression of repressors might be very important is when strong direct negative feedback loop exist for activators. Such repressors might exist as miRNAs (Kuypers et al. 2016; Zhao et al. 2010) or as other transcription factors (Pozniak et al. 2010). Repressing such repressive loops might not only help activate some genes which are hard to activate, but also increase the level of activation in general by removing endogenous constraints. Lastly, some repressors might dominate activators indirectly by repressing the chromatin or regulating RNA metabolism. In such a situation, repressing these chromatin modulating genes might help activators achieve their function faster.

To summarise, if sTFs are to become potent tools for cell type programming a number of key issues need to be addressed. We need to learn how to robustly activate or repress any gene of interest, but also how to achieve sustained high levels of activation to raise above transcriptional noise. Moreover, if sTFs are to become highly efficient for cell fate programming, we need to learn how to modulate large number of genes at the same time in the same cell. Exploring these issues is the focus of this thesis, in the context of neural lineages.

1.4 Neural stem cells and oligodendrocyte differentiation

1.4.1 Neural stem cells can differentiate into oligodendrocytes, astrocytes and neurons

For most of the 20th century the existence of neural stem cells (NSCs) in the adult brain was thought unlikely. During the 1990s and early 21st century, however, it has been demonstrated that neural stem cells can be isolated from certain adult brain regions, and there are likely two zones of neurogenesis that are conserved in mammals (Reynolds & Weiss 1992): including hippocampus and sub-ventricular zone (Eriksson et al. 1998; Lois & Alvarez-Buylla 1993). Importantly, nowadays we are able to capture and maintain these stem cells in culture (either as heterogeneous neurospheres or as adherent monolayer) for prolonged periods of time (Conti et al. 2005). After injury NSCs proliferate and may contribute to the repair of damaged regions (Arvidsson et al. 2002). Stimulating endogenous repair is therefore an ongoing area of research.

To be able to produce new neurons throughout adult life, NSCs must balance self-renewal (proliferation) and differentiation. NSCs often exist in a quiescence or dormant state. After stimulus from the environment, quiescent NSCs re-enter cell cycle, proliferate and generate new neurons and glia. Each stage of this process is closely regulated by extrinsic signals and intrinsic regulatory networks; known regulatory pathways include: Wnt/ β -catenin, Notch and Shh (Morrison et al. 1997; Ahmed et al. 2009). The balance between these pathways and how dormancy, quiescence, activation and proliferation are controlled in the endogenous NSCs is still

unclear and is an area being explored by many groups. In vitro, NSCs are expanded in EGF and FGF-2 driving them into a highly proliferative 'active' state.

NSCs also have to be able to differentiate to generate neurons, astrocytes and oligodendrocytes. Knowledge of the process and pathways involved in differentiation have been obtained from many decades of research in developmental biology. NSCs self-renew in presence of EGF and FGF-2 (Reynolds & Weiss 1992). EGF withdrawal from the culture media can result in differentiation to all three standard CNS cell types. Moreover, it is possible to induce differentiation to specific lineages by exposure to LIF or BMP (Bonaguidi et al. 2005) to generate astrocytes; alternatively, FGF-2, PDGFR-AA and Forksln directs NSC towards oligodendrocytes. These have become standard differentiation regimes for neural stem cell assays in vitro and are used in this thesis. However, a major issue is variability between lines and for oligodendrocyte differentiation a lack of efficient differentiation. Thus, improved methods to control NSC differentiation are needed.

A number of different markers are used to monitor differentiation state. Nestin, Sox1 and Sox2 are often used as NSC markers; β -tubulin as a neuronal marker (detected using the antibody TuJ1); the intermediate filament protein GFAP is an astrocyte marker; oligodendrocyte progenitors are marked by Sox10, PDGFR α or O4 (early markers), while terminally differentiated oligodendrocytes express myelin basic protein (MBP) (Conti et al. 2005).

1.4.2 Genetic profile of oligodendrocytes

Oligodendrocytes (OLs) are one of the key cell types produced by NSC differentiation. Their role is to ensure efficient and fast propagation of action potential

in central nervous system by myelinating neurons. OLs arise from precursor cells known as oligodendrocyte precursor cells (OPCs) which themselves arise from sub-ventricular cells in the brain (Emery 2010). Differentiation of OPCs to OLs *in vivo* is under strict temporal and spatial regulation and also goes through an intermediate stage of OLs which are premyelinating (Richardson et al. 2006) (Figure 1-7). Loss of myelinating OLs results in neuron degradation as seen in demyelinating diseases, such as multiple sclerosis and various leukodystrophies. Obtaining a consistent large-scale *ex vivo* source of OPCs and OLs would potentially open new ways to treat these diseases using cell-based therapies (Givogri et al. 2006). To better understand how we can generate these cell types, we need to better understand mechanisms of lineage commitment and differentiation in OPCs and OLs.

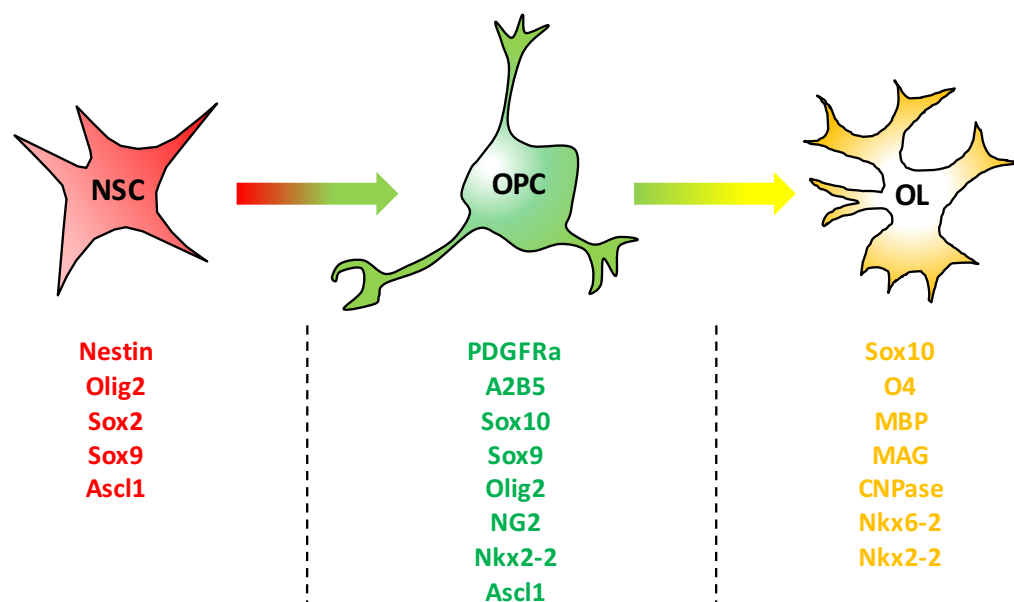


Figure 1-7 | Generic summary of key markers and regulatory proteins during OPC specification and subsequent oligodendrocyte differentiation.

Gene	Gene function	Reference
Sox10	Involved in the development of neural crest; specific to oligodendrocyte in CNS glial cells.	Najm et al. 2013 Christine D Pozniak et al. 2010
Olig2	Protein interacts with Sox10 and Nkx2-2; involved in oligodendrocyte differentiation	Yang et al. 2013 Maire et al 2009
Nkx2-2	Repress neuron differentiation; Involved in oligodendrocyte development; involved in astrocyte differentiation	Wang et al. 2014 Najm et al. 2011a
Nkx6-2	Involved in oligodendrocyte differentiation; role in CNS myelination	Cai et al. 2005 Cai et al. 2010
Ascl1	Involved in neurogenesis; involved in oligodendrocyte differentiation.	Braun et al. 2015 A. Raposo et al. 2015

Table 1-2 | Key transcription factor ‘master regulators’ involved in the oligodendrocyte lineage.

Some of the key genes involved in OPCs include: *Sox10*, *Olig2*, *Nkx2-2* and *Ascl1* (Table 1-2). This has been explored by investigating ESC and iPS cell differentiation and key events in mouse development. During transition from neural rosettes (early neuroprogenitors derived from ESCs) to OPCs, *Pax6*, *Sox9* and *Neurog1* is downregulated, *Ascl1* remains at similar levels, while *Nkx2.2*, *Sox10*, *Sox17*, *Olig2*, *Olig1* and *PDGFR α* are up-regulated (Emery 2010; Wegner 2008; Najm et al. 2011b). During differentiation of OPCs to premyelinating OLs and then to myelinating oligodendrocytes, MBP, Plp1 and Mag are upregulated, while expression of *Sox10*, *Olig2*, *Olig1*, *Ascl1* and *Nkx2.2* are maintained in postmitotic OLs (Emery 2010; Wegner 2008). Knock-out of either *Olig2* (Lu et al. 2002) or *Sox10* has been shown to result in loss of oligodendrocytes. Moreover, loss of myelin gene regulatory factor (MRF) has been linked to impaired OPC-OL transition (Emery et al. 2009), while other genes, like *Sox6*, *Hes5*, *Id4* and *Id2* have been shown to play an active role in preventing OPCs from differentiating further (Emery 2010). Together, these transcription factors and associated gene regulatory networks define OPC and oligodendrocyte identity (Fig. 1-7).

1.4.3 Methods for generation of oligodendrocytes ex vivo

Multiple methods have been developed to allow OPC and OL production from mouse or human foetal NSCs – often trying to mirror events used in development in the culture dish. The most basic approach is exposing NSCs to FGF and PDGF-A which results in about 20% of cells becoming PDGFR α positive OPCs (Glaser et al. 2007). This protocol has been further optimized by a various groups (Franco et al. 2015; Neri et al. 2010; Wang et al. 2015). In a number of studies, ectopic overexpression of either *Sox10* (C. D. Pozniak et al. 2010a; Wang et al. 2014), *Olig2*

(Coprav et al. 2006; Maire et al. 2009) or both (Liu et al. 2007) has been shown to facilitate or even induce NSC differentiation toward oligodendrocyte lineage. However, this approach cannot provide immunologically matched cell types for cell therapy due to difficulty of obtaining adult NSCs.

Oligodendrocytes have also been successfully generated by directed differentiation of ESCs. A number of studies have defined protocols for deriving OLs from ESCs (Neman & de Vellis 2012; Najm et al. 2011b; Sundberg et al. 2010). These methods usually make use of complex growth medium and can last for up to 48 days to make mature OLs. Moreover, it is relatively hard to calculate efficiency of these protocols accurately as unknown number of cells die and proliferate during the differentiation protocol.

Similarly to OL differentiation from ESCs, methods has been developed for turning patient iPSCs to oligodendrocytes (Wang et al. 2013; Douvaras et al. 2014). As these cells are taken through normal developmental process, protocols are very lengthy often lasting for over 100 days to make OLs (Wang et al. 2013). The process involves a number of different growth mediums which have to be added at specific time points (Wang et al. 2013). Although these methods have a lot of promise for therapeutic potential, they are very lengthy and labour-intensive. Thus, it is of the utmost importance to find new ways to bypass the inherent variability, lack of efficiency, long time-scales and experimental difficulties of current pluripotent and multipotent stem cell differentiation protocols. One interesting aspect is forward programming of iPSC. This involved overexpression of key lineage master regulators (*Sox10* in case of OPC/OL lineage) and then inducing differentiation. Such approach has been reported recently to be highly efficient in driving OPC production from

hiPSC. In the study by Pawlowski and colleagues, overexpression of *SOX10* and *OLIG2* gave rise to CNPase and PLP positive cells in 20 days (Pawlowski et al. 2017).

One alternative is transdifferentiation. This has been reported for mouse embryonic fibroblasts to OPCs and OLs (Najm et al. 2013; Yang et al. 2013). In the methods described for NSC and iPSC, scientists are usually making use of growth factors to generate OPCs and OLs. However, to achieve transdifferentiation, overexpression of specific transcription factors is used – and is more reminiscent of the ‘Yamanaka’ type approach. This is often called direct reprogramming. In these studies, overexpression of three cell type specific transcription factors (Sox10, Olig2 and Nkx6-2 or Sox10, Olig2 and Zfp536) was sufficient to achieve direct cell type conversion. The resulting OLs had markers like PLP1, O4, MBP and were capable of making myelin sheets after transplantation. The efficiency, however, remains relatively low, with less than 1% of original cells becoming OPCs.

The major focus of this thesis is to test if sTFs can enhance NSC differentiation to OPC/OLs and be used to induce efficient MEF transdifferentiation to OPC/OLs. i.e Can sTFs be useful in programming NS cells or reprogramming MEFs for production of OPCs/OLs?

1.5 Aims and hypothesis

We hypothesize that sTFs when delivered in pools of activators and/or repressors will be capable of transcriptional resetting to drive cell type lineage conversions.

In particular, we have focused on the following specific questions:

1. Can we activate expression of endogenous *Sox10*, *Olig2* and *Nkx6-2* in mouse neural stem cells and embryonic fibroblasts using sTFs?
2. Can we develop an improved 'toolkit' for delivery of multiple sTFs?
3. Can we simultaneously activate and repress distinct target genes using sTFs?
4. Can we steer differentiation of NSC to OPC using sTFs?
5. Can we trigger transdifferentiation of MEFs to OPC by sTFs?

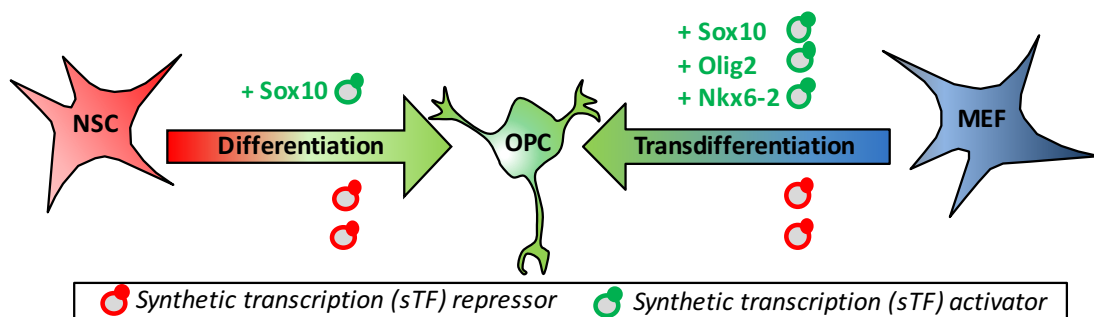


Figure 1-8 | Graphical representation of main hypothesis

2 Methods and Materials

2.1 Cell culture

2.1.1 Neural stem cell culture medium

Mouse neural stem cells (NSCs) were cultured using self-renewal medium (DMEM/F-12 supplemented with Glucose, Non-Essential Amino Acids (NEAA), Pen/Strep, HEPES, 75 mg/ml BSA solution, 2-mercapthoethanol, B27 and N2 supplements, 1ng/ml EGF, 10ng/ml FGF, 1 μ g/ml laminin) as previously reported (Conti et al. 2005). Cells were split using Accutase (Sigma, #A6964) every 2-3 days or when reaching ~80% confluence. Typically, cells were used around 10-30 passages.

2.1.2 Mouse embryonic fibroblast cultures

Mouse embryonic fibroblasts (MEFs) lines were established from E13.5 mouse limbs. Limbs from 13.5 days old mouse embryos were dissected and incubated in 0.25% Trypsin/EDTA for 30 min at 37°C. The cell suspension was triturated using a P1000 to disrupt tissue and dissociate cells. The cells were passed through a cell strainer prior to plating in flasks (Greiner Bio-one, #542042). Harvested cells were incubated in MEF culture medium (see below).

2.1.3 MEF culture medium

MEFs were cultured using AmnioMAX C-100 Basal medium (Thermo Fisher Scientific, #17001-074) with AmnioMAX C-100 supplement (Thermo Fisher Scientific, #12556-023). MEFs were passaged every 2-3 days. 0.25% Tripsin/EDTA was used to harvest cells from plates.

2.1.4 HEK293T culture medium

Human embryonic kidney (HEK293T) cells were incubated in DMEM medium containing 10% Fetal Bovine Serum (Life Technologies, #10270106), 0.045g/ml penicillin/streptomycin (Life Technologies, #15140122), 100x non-essential amino acids (Life Technologies, #11140050), 1mM Sodium Pyruvate Glutamate (Life Technologies, #11360088) and 2mM L-Glutamine (Life Technologies, #25030123). HEK293T cells were grown in 37°C incubator and split every 2-3 days using Accutase.

2.1.5 Freezing and thawing cell lines

For short term cryopreservation of cells, they were stored in -80°C (2-3 months). Liquid nitrogen was used for longer archiving. Cells were frozen down by resuspension in DMEM containing 10% DMSO and immediately transferred to -80°C. To thaw the cells a cryo-vial was rapidly warmed in a 37°C (water bath) and 10ml of warm growth medium gently added to the cell suspension. The resulting solution was centrifuged to remove DMSO and debris and cells were resuspended in an appropriate volume of warm growth medium on tissue culture plastic.

2.1.6 NSCs differentiation medium and protocol

NSCs were triggered to differentiate to OPCs using following differentiation protocols. The day prior to differentiation the cells were seeded at 2×10^4 cells per each well of a 6-well plate. The next day, fresh culture media lacking EGF but containing 10 μ M Forskolin (Cambridge BioScience; #11018) and 10ng/ml PDGF-AA (R&D System, #221-AA-010) and FGF-2 (10ng/ml) was added. Cells were left in such

differentiation medium for four days. After four days, cells were either analysed or induced to differentiate to OLs using medium described below.

To differentiate OPCs to OLs, the following medium was used. Medium lacking all growth factors, but containing 30ng/ml T3 (Sigma; #T6397) and 200 μ M L-Ascorbic acid (Sigma; #A4544) was added to cells. Cells were analysed after four days.

2.1.7 MEF transdifferentiation to OPCs

We attempted transdifferentiating MEFs to OPCs using the following protocol: cells were transfected (see below) and left in MEF culture medium for three days. After three days, the MEF culture media was exchanged with MEF transdifferentiation (TD) media. TD media contained following: DMEM/F12 (Invitrogen, #11320) supplemented with 1:100 N-2 (LifeTech; #17502-048), 1:50 B-27 (LifeTech; #17504-044), 2mM L-Glutamine, 200ng/ml SHH (R&D Systems; #461-SH-025/CF), 20ng/ml FGF2 (R&D Systems), 20ng/ml PDGF-AA (R&D Systems) and 2 μ g/ml Laminin (Sigma). This media should be permissive for OPCs.

To force transdifferentiating MEFs to differentiate to OL, after 14-21 days in TD media the cells were incubated in the following medium for 3 days: DMEM/F12 (Invitrogen, #11320) supplemented 1:100 N-2 (LifeTech; #17502-048), 1:50 B-27 (LifeTech; #17504-044), 2mM L-Glutamine, 40 ng/ml T3 (Sigma), 200ng/ml SHH (R&D Systems; #461-SH-025/CF), 200 μ M L-Ascorbic acid (Sigma; #A4544). This is adapted protocol based on previously published by Najm *et al.* 2013.

2.1.8 Clonal cell line derivation

To derive clonal cell lines, 500-1000 cells were seeded into 10cm² dish and left for 10-14 days. Colonies that grew were manually picked using a P200 and colony picking tips, and re-plated into each well of a 96-well plate for further expansion. After sufficient cells expansion, they were analysed either by genotyping, qPCR or immunostaining to validate positive colonies.

2.2 Mammalian cell transfection (MEF, NS and HEK293T)

2.2.1 Polyethylenimine (PEI)

HEK293T cells were transfected using the PEI transfection reagent and the following method. A day before transfection, 5x10⁴ cells were seeded into 24-well plate. On the day of transfection, 500ng of DNA and 2μl of PEI was added to 75μl DMEM (no supplements), mixed and incubated at room temperature for 15 min. Meanwhile, 1ml of pre-warmed culture medium was been added to cells. DNA/PEI/DMEM solution was then added to cells dropwise. Medium was exchanged 18h post-transfection.

2.2.2 Nucleofection

NSCs were transfected using Amaxa Nucleofector 4D system (Lonza, #V4XC-3032). Unless stated otherwise, 2x10⁵ cells were resuspended with P3 or SG Solution transfection buffer (depending on the NSC line). 400ng of DNA was added to the solution and cells were transfected using the DS-113 or DN-100 transfection protocol using 16-well transfections strips. 80μl of culture medium was added to each well of

strip and all content of each well was moved to either 6-well plate or 12-well plate. An appropriate volume of relevant culture medium was added to each well.

2.2.3 Lipofection

MEFs were transfected using Lipofectamine 3000. Cells were seeded a day prior to transfection into a 6-well plate. On the day of transfections, 2 μ g of DNA was mixed with 250 μ l OptiMEM medium (Thermo Fisher Scientific; #31985062), 5 μ l Lipofectamine 3000 Reagent and 5 μ l P3000 reagent (Life Technologies; #L3000001). The DNA/Lipofectamine solution was incubated at room temperature for 15 min. After incubation, the DNA/Lipofectamine mix was added to cells.

2.3 Immunocytochemistry

Cells were washed with PBS and fixed with 4% PFA for 10 min at room temperature. After fixation, cells were washed three times with PBS. Cells were then blocked at least for 1 hour in blocking solution (1%BSA, 3%goat serum, 0.1%Triton in PBS unless stated otherwise). The blocking solution was removed and the primary antibody was added and incubated at 4°C overnight in blocking solution (see Table 2-1). After staining with primary antibody, cells were washed three times with PBSt for 15 min. Cells were then stained with an appropriate secondary antibody (1:1000 dilutions) for at least 1 hour in the dark at room temperature. Cells were washed two times with PBS, stained with DAPI (1:5000 in PBS) for at least 5 min and washed with PBS twice.

Antibody	Specie	Dilution	Supplier, cat. number
OLIG2	Rabbit	1:400	EMD Millipore, #Ab9610
GFP	Mouse IgG1	1:400	Sigma, #G6539
V5	Mouse IgG2b	1:1000	eBioscience, #14-6796-82
SOX2	Rabbit	1:100	Abcam, #Ab92494
SOX10	IgG1	1:10	Hybridoma
SOX10	IgG1	5 μ g/ml	R&D System, #MAB2836
SOX10	Goat IgG	1:100	R&D System, #AF2864
SOX10	Rabbit	1:100	Abcam, #Ab27655
SOX10	Rabbit	1:100	Brown's lab
SOX10	Rabbit	1:200	Abcam, #Ab155279
HA	Mouse IgG1	1:100	Cell signalling, #2367
O4	Mouse IgM	1:200	Charles ffrench-Constant lab
NG2	Rabbit	200	Abcam, #Ab83178
MBP	Rat	1:250	Abcam, #Ab7349

Table 2-1 | Summary of primary antibodies used.

2.4 Microscopy and image processing

Fluorescent microscopy images were taken using either the Operetta or Nikon TiE wide-field fluorescent microscopes. Images were processed using either the NIS-Elements or Columbus software.

2.5 Image cytometry

Image cytometry was being used to quantify immunostaining results. Stained cells were screened using Operetta high content screening platform (Perkin Elmer), processed on Columbus software and analysed using TIBCO SpotFire software. 'Find Nucleus' protocols were unique each time as it strongly depended on the cell type, number of cells and confluence in the well as well as shape of the nucleus. However, protocol C (see Columbus software) was used as a starting point in most cases. Gating and signal thresholds were set using negative controls (aka non-treated cells).

2.6 Flow cytometry

Cells were analysed using LSR Fortessa 5 or 4-laster flow cytometer (BD BioScience). Unless stated otherwise, 10,000 events were collected using forward scatter threshold of 5000. Fluorescence data was collected using following cytometer settings: 488 nm laser and B530/30-A nm bandpass filter for GFP/NeoGreen, 561 nm laser and YG586/15-A nm bandpass filter for mRuby2/ mCherry, 405 nm laser and V450/50-A nm bandpass filter for DAPI. Data was analysed using the FlowJo software.

When cell surface markers were assessed, we immunostained using the following protocol: cells were lifted with Accutase and incubated with PBS containing 2% FCS, 1:200 Fc block CD16/CD33 (BD Pharmingen, #553141) and appropriate primary antibody. Cells were left for 30 min at room temperature. After incubation, the cells were washed with 2% FCS/PBS and incubated with appropriately diluted (in 2% FCS/PBS) secondary antibody (usually 1:1000). Cells were left for 30 min at room temperature in dark. After incubation, cells were washed with 2% FCS/PBS and analysed with the flow cytometer. Stained but non-treated (no transfection or differentiation) cells were used as negative control for gating.

2.7 TaqMan quantitative RT-PCR

2.7.1 RNA extraction

RNA was extracted using QIAGEN RNeasy Mini Kit (Qiagen; #74104) according to the manufactures protocol. RNA was stored at -80°C.

2.7.2 Reverse transcriptase (RT) reaction

To generate cDNA from RNA, the reverse transcriptase reaction was conducted using the SuperScript II Reverse Transcriptase Kit (Life Tech; #18064014). The detailed protocol can be found on suppliers' website. 1-2ug of RNA was typically used in each RT reaction. cDNA was stored at -20°C.

2.7.3 qPCR

qPCR was performed using TaqMan Fast PCR Master Mix (Thermo Fisher; #4352042) and TaqMan assays (Table 2-2). Each reaction had a final volume of 10 μ l: 5 μ l PCR Master Mix, 0.5 μ l specific TaqMan Probe, 2 μ l cDNA, 2.5 μ l water. Each sample was run as a technical duplicate.

Target	Assay	Specie
<i>Gapdh</i>	Mm99999915_g1	Mouse
<i>Sox10</i>	Mm00569909_m1	Mouse
<i>Olig2</i>	Mm01210556_m1	Mouse
<i>Nkx6-2</i>	Mm00807812_g1	Mouse
<i>PDGFRα</i>	Mm00440701_m1	Mouse
<i>Ascl1</i>	Mm04207567_g1	Mouse
<i>Sox2</i>	Mm03053810_s1	Mouse
<i>Sox9</i>	Mm00448840_m1	Mouse
<i>Nkx2-2</i>	Mm00839794_m1	Mouse
<i>GAPDH</i>	Hs99999905_m1	Human
<i>SOX2</i>	Hs00366918_m1	Human
<i>OLIG2</i>	Hs00300164_s1	Human
<i>SOX10</i>	Hs00366918_m1	Human

Table 2-2 | List of TaqMan assays used.

2.8 RNA-FISH

For coverslip coating each coverslip was placed in a 6-well plate and incubated in 10 μ g/ml poly D-Lysine (PDL) in sterile water for 5 min at room temperature. PDL was removed and the coverslip was incubated in 20 μ g/ml laminin in DMEM overnight in 37°C. Cells were seeded onto the coverslip and maintained in the incubator until staining. Cells were washed with PBS and then fixed with 37% formaldehyde solution for 10 min in room temperature. Cells were permeabilized using 70% ethanol for at least for 1 hour at 4°C. Cells were washed with buffer A (BioSearch Technologies; #SMF-WA1-60) for 5 min at room temperature. Samples then were stained using hybridization buffer (BioSearch Technologies; #SMF-HB1-10) and RNA-FISH probe mixture (custom made probes or Mouse GAPDH; BioSearch Technologies; #SMF-3002-1) overnight at 37°C in dark. Cells were washed with buffer A for 30 min in dark at 37°C and then stained with DAPI (1:5000) for 30 min in dark at 37°C. The coverslip was washed with buffer B (BioSearch Technologies; #SMF-WB1-20) for 5 min at room temperature. The coverslip was mounted onto the slide together with ProLong Gold Antifade Reagent (New England BioLabs; #9071S) and imaged on a Nikon TiE microscope with 60x oil objective.

2.9 Molecular cloning

2.9.1 Site-directed mutagenesis

Site-directed mutagenesis was performed using New England BioLabs Q5 Site-Directed Mutagenesis kit (New England BioLabs; #E0554S). The suppliers' guidelines were followed. For PCR primer sequence see Appendix.

2.9.2 Gateway cloning

All gateway cloning was performed using Invitrogen Gateway BP or LR Clonase II Enzyme kits (Invitrogen; #11789020 and #11791020) and following suppliers' instructions.

2.9.3 Gibson cloning

For Gibson reactions, an in-house Gibson Mix was used. This mix contained: 5x isothermal buffer (0.5M Tris-HCL (Sigma-Aldrich), 50mM MgCl₂, 1mM dNTPs (Thermo Fisher Scientific), 50mM DTT (Life Tech), 0.25g/ml PEGC8000 (Sigma-Aldrich), 5mM NAD (Sigma-Aldrich)), T5 exonuclease (10U/ μ l), Phusion DNA polymerase (2U/ μ l) and Taq DNA ligase (40U/ μ l).

15 μ l of Gibson mix (described above) was mixed with 100ng linearized backbone vector and equimolar amounts of other assembly fragments. Reaction mixes were incubated at 50°C for 60 min. 2 μ l of the reaction mix was transformed with competent *E.coli*.

DNA fragment overlaps were usually between 16-40 base pairs (over 25 base pairs for more complicated reactions). Overhangs were designed to have 72°C melting temperature (T_m). Incubation times were sometimes increased or lowered to avoid incorrect assembly if complicated constructs were being generated.

2.9.4 Golden Gate Assembly

All Golden Gate (GG) reactions have been performed using following method: GG enzyme master mix (10 μ l T4 Ligase buffer, 1 μ l 100xBSA, 5 μ l appropriate restriction enzyme, 10 μ l T4 ligase) was prepared in bulk and stored in -20°C. To

perform a GG reaction, 2.6 μ l of the enzyme master mix was mixed with 50ng empty entry plasmid, 50ng of each part plasmid and topped up with water to 10 μ l. The following program was run on a Thermocycler to start the reaction: 37°C for 5 min; 15-20 cycles of 37°C for 5 min and 16°C for 10 min; 5 min at 50°C; 10 min at 80°C; stored at 4°C. Home-made competent *E.coli* bacteria were transformed with 5 μ l of GG reaction mix.

2.9.5 *E. coli* transformation

To propagate plasmids, competent *E.coli* were transformed using following protocol. 25 μ l of competent *E.coli* bacteria were mixed with 2-5 μ l of plasmid/reaction solution (usually around 1-100ng of plasmid) and kept on ice for 30 min. The bacterial/plasmid solution was incubated at 37°C for 30s and then put back to ice for 2 min. 300-500 μ l of LB (no selection) was added and mixture was incubated at 37°C shaker for 60 min. Transformed bacteria were either plated out on LB agar plate (containing appropriate selection) or mixed with larger volume of LB medium (containing appropriate selection) and incubated overnight at 37°C incubator or shaker, respectively.

Plasmid ID	Plasmid name	Source
M1	TRE_SOX10_2A_OLIG2	Cloning
M19	CAG_mCherry	Ian Chambers' lab
M28	CAG_dCas9_p300	Patric Kai's lab
M29	CAG_dCas9_KRAB	Patric Kai's lab
M33	CAG_dCas9_VPR	AddGene #63798
M34	CAG_VRER_dCas9_KRAB	Cloning
M35	CAG_VRER_dCas9_VP160	Cloning

M39	Pb_dCas9_VP160_mCherry (Empty Ai1)	Cloning
M51	CAG_dCas9_SON	Cloning
M52	CAG_dCas9_tSON	Cloning
M53	CAG_dCas9_S10	Cloning
M54	CAG_dCas9_VP160	AddGene #48227
M55	CAG_dCas9_2SON	Cloning
M56	CAG_dCas9_3SON	Cloning
M59	CAG_dCas9_S10S9	Cloning
M60	CAG_dCas9_S1	Cloning
M61	CAG_dCas9_S2	Cloning
M62	CAG_dCas9_S5	Cloning
M63	CAG_dCas9_S6	Cloning
M64	CAG_dCas9_S7	Cloning
M65	CAG_mNG (Plasmid 1)	Patric Kai's lab
M66	CAG_mNG_2A_mR2 (Plasmid 2)	Patric Kai's lab
M67	CAG_dCas9_TAD1	Patric Kai's lab
M68	CAG_dCas9_TAD2	Patric Kai's lab
M69	CAG_dCas9_TAD3	Patric Kai's lab
M70	CAG_dCas9_TAD4	Patric Kai's lab
M71	CAG_dCas9_TAD5	Patric Kai's lab
M72	CAG_dCas9_TAD6	Patric Kai's lab
M73	CAG_dCas9_TAD7	Patric Kai's lab
M74	CAG_dCas9_TAD8	Patric Kai's lab
M75	U6-BsaI-T3	Mali <i>et al</i> 2013
M76	CAG_GFP	pmaxCloning™ Lonza
M77	CAG_pBase	Steve Pollard's lab
M78	CAG_TET3G	Steve Pollard's lab
M78	CAG_Sox10_2A_Olig2	Cloning

Table 2-3 | List of vectors used in study.

2.10 Design and production of U6-gRNA expression plasmids

gRNAs were designed using CHOPCHOP web tool (<https://chopchop.rc.fas.harvard.edu/index.php>) and appropriate overhangs for ligation were added to sequences. These sequences were ordered via Sigma Oligo Synthesis service. gRNA expressing plasmids were constructed using the following protocol: Synthesized oligos were phosphorylated with Fermentas T4 PNK (Thermo Fisher Scientific; # EK0031). In parallel, the backbone vector for gRNAs (SP149) was prepared by digestion with the BsaI restriction enzyme, cleaned up with MinElute PCR purification kit (Qiagen; #28004) and dephosphorylated with SAP enzyme (New England BioLabs; #M0371). To ligate the phosphorylated gRNA oligos with dephosphorylated backbone, T4 DNA Ligase (New England BioLabs; #M0202) was used. The resulting plasmids were transformed and a small amount digested with EcoRI and BsrGI restriction enzymes to confirm correct assembly.

2.11 Gene tagging with V5 using single-strand oligonucleotide

To knock in a V5 tag to create a *Sox10-V5* fusion product, we designed guides to cut near the STOP codon. Recombinant Cas9 protein was used. The following protocol was used to deliver reagents: 0.8 μ l of crRNA (250 picomoles/ μ l) was mixed with 0.8 μ l of tracrRNA (250 picomoles/ μ l) and 3.4 μ l water. The mix was incubated at 95°C for 5 min and then 30 min at room temperature. 2.5 μ l of above mix was mixed with 10 μ g purified CAS9 protein and left for 10 min at room temperature. The above solution was then mixed with 0.6 μ l ssDNA (50 picomoles/ μ l stock) and transfected into NSC cells using DN-100 Nucleofector protocol.

To test if tagging was successful, genomic DNA of the transfected cells was harvested using DNeasy extraction kit (Qiagen; #69504). PCR reactions were performed using harvested genomic DNA as template and specific PCR primers which bind V5 (reverse) and *Sox10* exon (forward) sequence. The PCR product was run on an agarose gel to identify the correct band size.

3 Regulating mouse gene expression using sTFs

3.1 Activating *Sox10*, *Olig2* and *Nkx6-2* in mouse NS cells and MEFs

3.1.1 Construction of U6-driven gRNA expression vectors

To activate gene transcription using dCas9, gRNAs need to be targeted to the proximal promoter of the target gene of interest. More specifically, the 50bp to 450bp upstream of the TSS has been reported to be optimal for transcription activation (Gilbert et al. 2014). We therefore initially designed gRNAs within this region to target the three oligodendrocyte lineage regulators, *Sox10*, *Olig2* and *Nkx6-2* (Fig. 3-1 A).

Importantly, past studies have shown great variability in the levels of activation when using gRNAs. Thus, not every gRNA will be functional. We first designed ten distinct gRNAs spanning this region. ChopChop software has been used for this purpose (Fig. 3-1 B-D). gRNAs were selected based on their quality scores, taking into account GC content, self-complementary tendencies, off-targets, and their distribution across a range of positions within the promoter region.

For testing in mammalian cells the gRNAs were cloned into a mammalian expression plasmid (SP149). This contains a U6 promoter, gRNA scaffold and termination sequences and was originally reported in the study by Mali *et al* (Mali, Esvelt, et al. 2013). Cloning was achieved by linearizing the empty expression plasmid with BsaI-HF restriction enzyme, dephosphorylating this linearized backbone and ligating it with an annealed and phosphorylated oligos (Fig. 3-2 A). Successful plasmids were identified using test restriction enzyme digestions with BsrGI and EcoRI restriction enzymes (Fig. 3-2 B).

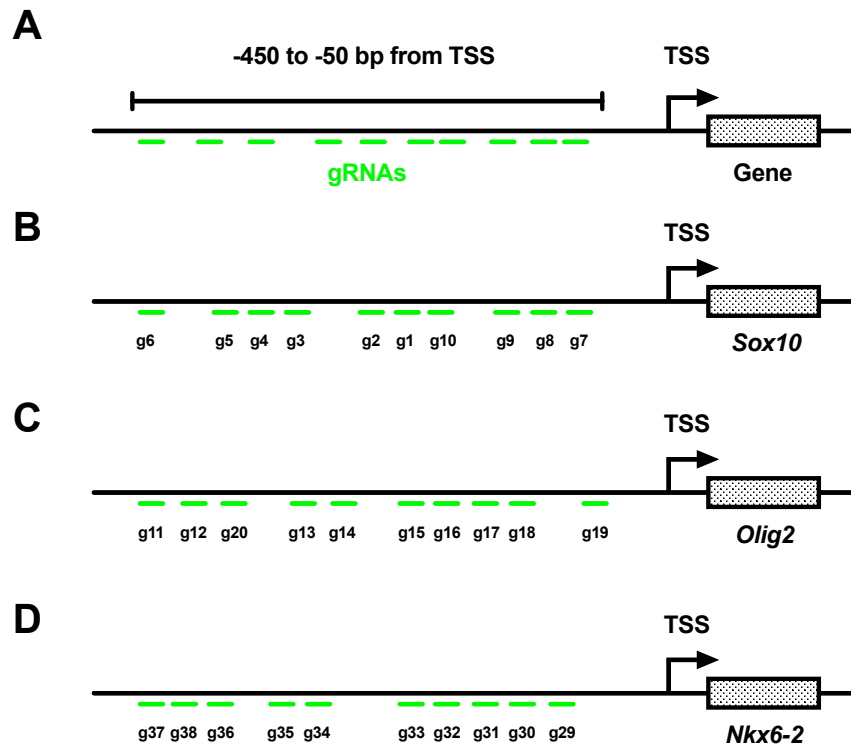


Figure 3-1 | gRNA placement for *Sox10*, *Olig2* and *Nkx6-2* gene activation.

(A) Strategy for design of gRNAs for transcriptional activation. (B, C and D) gRNA position along the proximal promoter and transcription start site (TSS) for transcription activation for *Sox10*, *Olig2* and *Nkx6.2*, respectively.

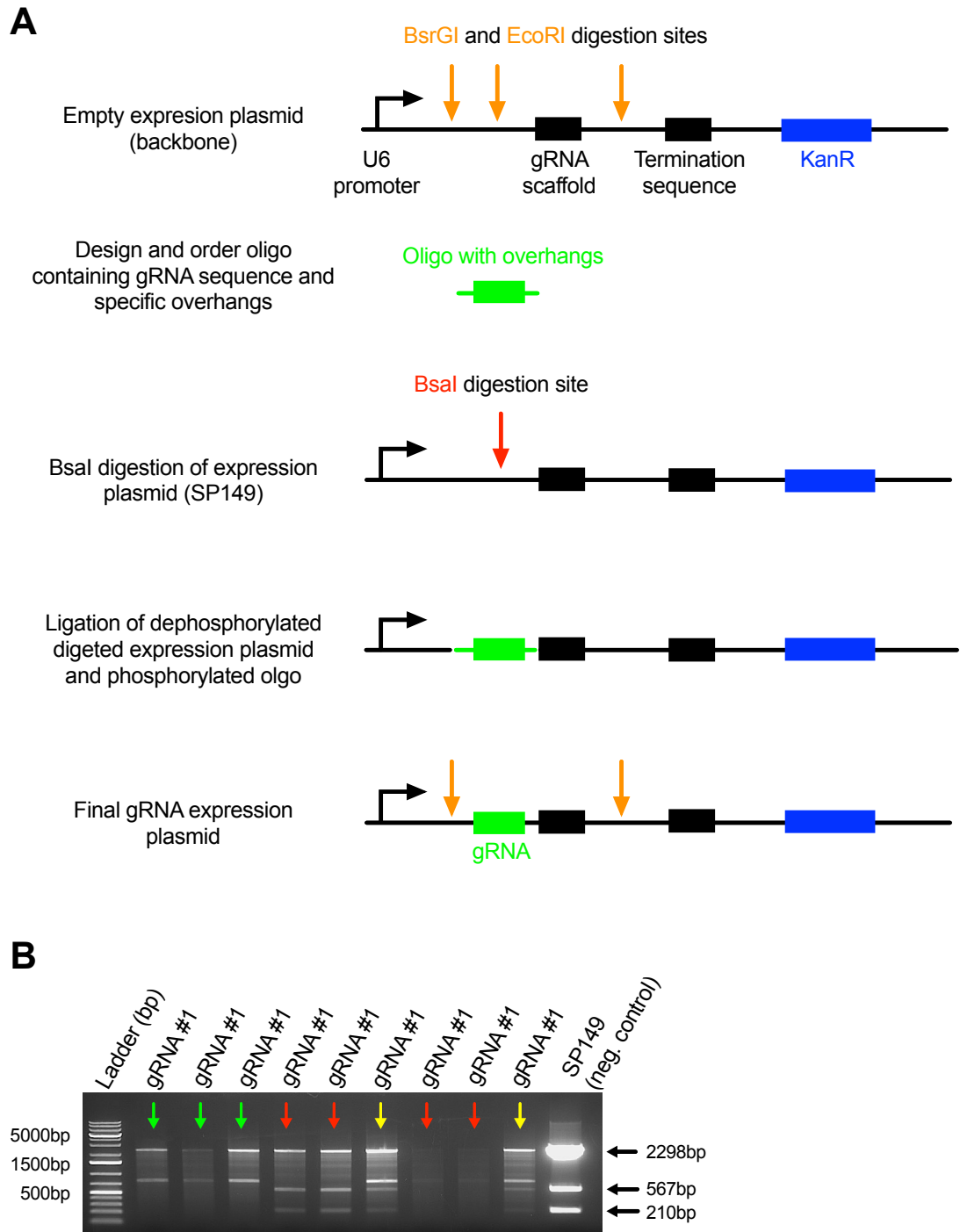


Figure 3-2 | gRNA-expression plasmid construction.

(A) Schematic representation of the process of cloning gRNAs into the U6 expression plasmid via digestion and ligation reactions. (B) Agarose gel electrophoresis to assess the restriction digestion product of cloned gRNA-expressing plasmids using BsrGI and EcoRI restriction enzyme; green arrows indicate correctly assembled plasmids, red arrows mean negative, yellow arrow indicate contaminated/mixed clones.

3.1.2 *Sox10*, *Olig2* and *Nkx6-2* activation in mouse NS cells and MEFs

To investigate if any of the designed gRNAs were functional when co-delivered with dCas9-VP160 in activating transcription of the target genes, we first assessed transfection of all ten gRNAs together using qPCR assays. This was carried out in both mouse neural (NS) cells and mouse embryonic fibroblasts (MEFs), in order to compare cell context effects. Based on the mRNA fold change, it was clear that *Sox10* can be activated in both cell types (>27.9-fold change in NSC and >14.1-fold change in MEF). By contrast for *Olig2* and *Nkx6-2* we could achieve increased expression only in the MEFs, (>184.8-fold change in mRNA for *Olig2* and >46.8-fold change for *Nkx6-2*) (Fig. 3-3 A).

Thus, some gRNAs which are validated as functional cannot work in other cell types, most likely due to cell specific biological restrictions in the transcriptional networks or epigenetic state of the target locus. Although *Nkx6-2* gene is not transcribed in either cell type, we find that only in MEFs can it be effectively activated (Fig. 3-3 A). This differing result means there must be some restriction in NS cells to the activity of gRNAs for *Nkx6-2*. This is unlikely to be due to differences in gRNA target sequences, as it would be unlikely that SNPs or other somatic mutations would be distributed across the TSS at so many distinct PAM sites. Possible reasons will be outlined in Discussion chapter.

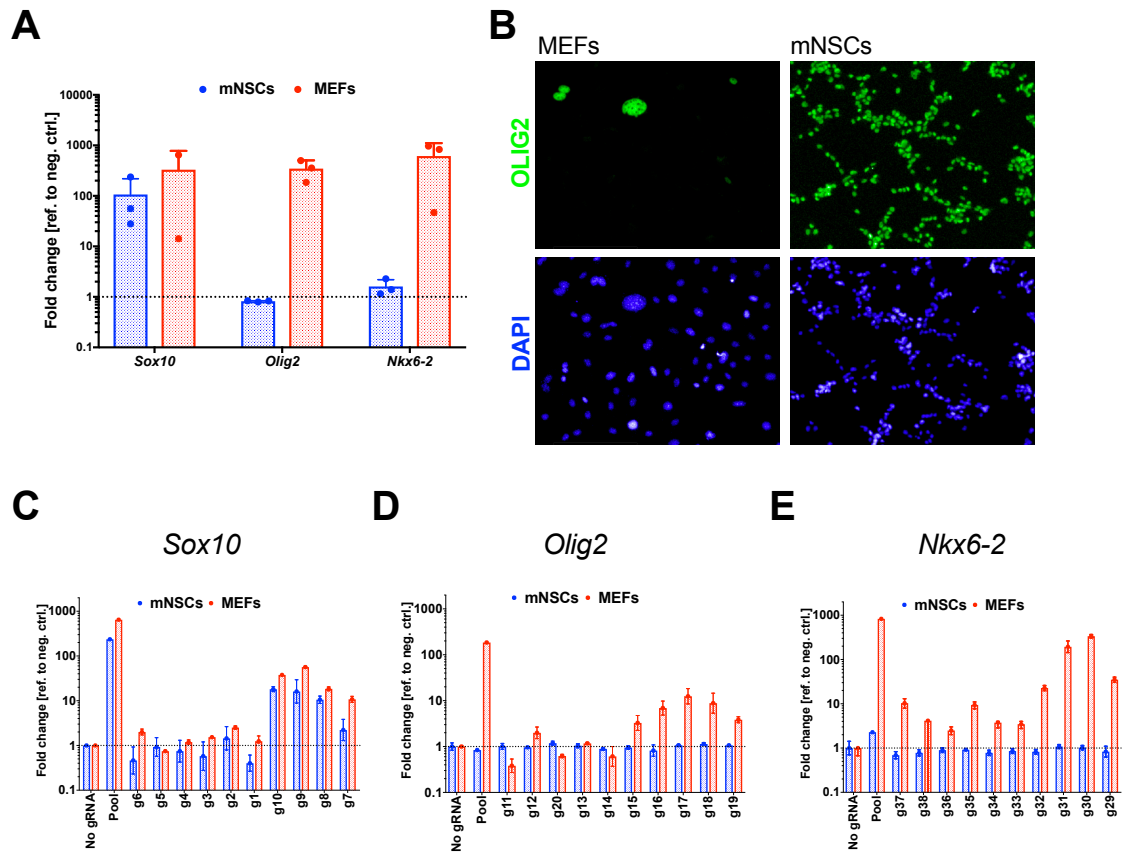


Figure 3-3 | *Sox10*, *Olig2* and *Nkx6-2* activation using sTFs in mouse neural stem cells (mNSCs; blue) and mouse embryonic fibroblasts (MEFs; red).

(A) qPCR results 3 days after transfection with dCas9-VP160 and pool of 10 different gRNAs for each of the target genes: *Sox10*, *Olig2* and *Nkx6-2*, two different cell lines, n=3. (B) Immunocytochemistry for *OLIG2* 3 days after transfection of mNSCs and MEFs with dCas9-VP160 and a pool of 10 gRNAs for *Olig2* (g11-g20). (C) qPCR results 3 days after transfection with dCas9-VP160 alone ("No gRNA") or dCas9-VP160 with either single gRNA ("g1", "g2", "g3" etc.) or a combination of 10 gRNAs ("Pool") targeting *Sox10*, n=1. (D) Single gRNA screen for *Olig2* (same setup as described for *Sox10* in C). (E) Single gRNA screen for *Nkx6-2* (same experimental setup as described for *Sox10* in C). Fold changes were normalized to the negative control (cells transfected with just dCas9-VP160).

Furthermore, the results from *Olig2* activation suggest that it is difficult to overexpress a gene that is already being transcribed. *Olig2* is high expressed by non-transfected NS cells, it is well established to be associated with the proliferative NS cell phenotype. However, it is fully repressed in MEFs. From activation experiments, it is evident that gRNAs targeting *Olig2* are functional as it can activate transcription in MEFs (Fig. 3-3 A-B). However, it shows no effect in NS cells, suggesting that if gene transcription is already high, over-activation is difficult. It seems easier to activate a non-expressed gene, then to overexpress an existing active gene using dCas9 sTFs. This observation has been reported in other studies as well (Fidanza et al. 2017; S. Lin et al. 2015).

To determine the relative activity of each individual gRNA MEFs and NSCs were transfected with dCas9-VP160 and each gRNA alone. The results reveal that from the pool of 10 gRNAs, 4 different gRNAs for *Nkx6-2*, 4 gRNAs for *Olig2* and 4 gRNAs for *Sox10* were effective in strongly activating transcription (Fig. 3-3 C-E). There was a trend towards increased activity of the gRNAs correlating to their proximity to the TSS. For each gene, the top 3 or 4 gRNAs (those that could activate >5-10 fold) were proximal to the TSS. This data also suggests that single gRNAs, even when functional, cannot achieve activation levels as high as the pool of 10 gRNAs. Consistent with the results from the pool of 10 gRNAs, individual gRNAs were unable to activate *Olig2* further from the high levels already present in NS cells. For *Nkx6-2* we could not activate transcription using single gRNAs at all in NS cells, despite MEF activation working well.

The greater levels of activation when using pools of gRNA can be explained by either synergistic or additive effects. To explore this further we next tested if co-

delivery of different combinations of the most active gRNAs is able to achieve same activation levels as a pool of 10 gRNAs (2 gRNAs were used for *Nkx6-2* and 3 for *Olig2*). This was again explored in both MEFs and NSCs. Cells were transfected (MEF – Lipofectamine; NSC – Nucleofection) with dCas9-VP160 and either a pool of 10 gRNAs or a pool of 2 for *Nkx6-2*, or 3 for *Olig2*.

For both *Nkx6-2* and *Olig2* gene activation in MEFs we found that a combination of 2-3 most potent gRNAs (g16-g18 in case of *Olig2* and g29-g30 in case of *Nkx6-2*) is able to achieve similar level of transcription activation as a pool of 10 gRNAs. However, large variations between fold-changes in each biological replicate was observed which made it hard to directly compare different experiments statistically (Fig. 3-4).

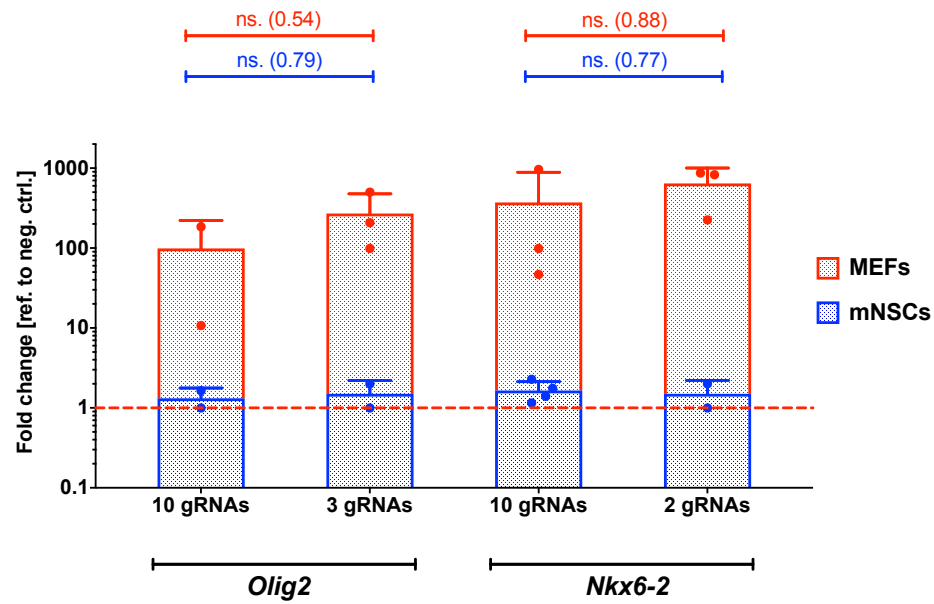


Figure 3-4 | Comparing small and large pools of gRNAs in their ability to activate *Olig2* and *Nkx6-2* in mouse NSCs and MEFs.

qPCR results 3 days after transfection with dCas9-VP160 and either combination of 2 (g29-g30 for *Nkx6-2*) or 3 gRNAs (g16-g18 for *Olig2*) or a combination of 10 gRNAs (g11-g20 for *Olig2*, g29-g38 for *Nkx6-2*), each dot represents biological replicate (for statistical comparison, two-tail t-test was done, p-values noted in the graph).

3.2 Exploring the synergistic versus additive effects of multiple gRNAs using single cell assays

The experiments in section 3.1 confirmed that dCas9-VP160/gRNAs can effectively activate transcription in genes of interest, and that effective design works best when gRNAs are proximal to the TSS. However, there are clearly cell type differences in the ability of these to activate. Finally, the data suggest that combinations of two or more gRNAs can work better than individual gRNAs.

qRT-PCR assays provide an average level of expression across the population. Therefore, it remains difficult to determine using this assay if only a few cells are activated to high level or all cells in the population to a lower level. The correlation between amount of sTF and target gene activation is therefore difficult to determine. To better understand whether there is an optimal level of expression and if sTFs work additively or synergistically we used immunocytochemistry and quantitative image cytometry.

3.2.1 Optimising a quantitative immunocytochemistry ('image cytometry') assay

Immunocytochemistry can provide single cell level data regarding relative levels of protein. To quantify the intensity of immunopositive cells we explored image cytometry tools that have become available with advances in automated microscopy and analysis software (Operetta, PerkinElmer). This technique is akin to flow cytometry in the type of data that is acquired. Using automated capture of large sets of microscopy images and subsequent image analysis many thousands of cells are scored at single cell level for levels of protein expression (Fig. 3-5 A). It removes many of

the biases of working with limited fields of view when using manual microscopy. A description of the image cytometry protocol can be found in the Materials and Methods chapter.

We first explored this methodology using mouse NSC. Cells were stained for DAPI and OLIG2, followed by analysis (Fig. 3-5 B). Analysis includes nucleus segregation and identification followed by collecting information of staining intensity. The cell nucleus identification and appropriate segregation is the most important part of analysis. Some cells (e.g. HEK293T cells due to clumping) require larger amount of effort to make identification possible while other cells (NSC and MEF) are less tricky as they grow more sparsely.

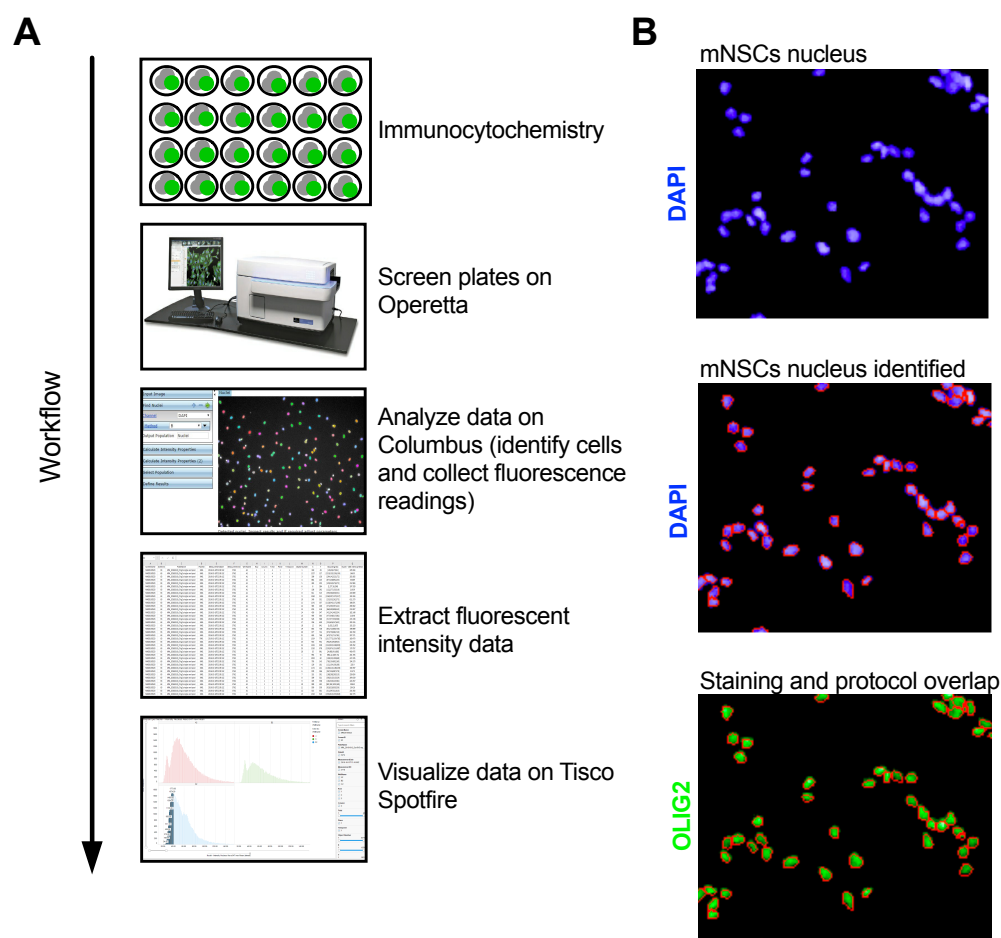


Figure 3-5 | Graphical representation of the image cytometry workflow.

(A) Cell culture plates (usually 24-well Corning® Costar® plates) are stained using immunocytochemistry and scanned on Operetta system to collect >10 images per well/condition. This usually cover at least 10,000 cells. Images are then analyzed using Columbus to identify and count cells (based on DAPI staining) and collect intensity measurements of fluorescence (which represents protein levels). Collected data is visualized and analyzed on Tisco Spotfire software. (B) Example image demonstrating one of the protocols and image analysis tool for identifying cell nuclei and then using identified area for measuring staining intensity.

3.2.2 Transcriptional activation of *OLIG2* in HEK293T using multiple gRNAs

We next tested if delivery of multiple gRNAs gives high activation due to a synergy between gRNAs or an additive effect (i.e. the sum of individual gRNA activities). For this purpose, we decided to use HEK293T cells and *OLIG2* as a target as these cells are easy to transfect and validated *OLIG2* antibodies exist. Five plasmids targeting *OLIG2* were constructed, containing either 3 gRNAs (g181, 182, and 183), 2 gRNAs (g181 and g182) or each individually. For the multiple gRNAs, these were cloned into a single vector with each driven by an individual U6 promoter (Fig. 3-6 A). These plasmids were generated using the Golden Gate assembly approach that is described in detail in Chapter 4; they were generated in alongside the work reported in that Chapter. It was important to include all the gRNAs on a single plasmid as co-transfection of separate plasmids is often unreliable.

HEK293T cells were transfected with these plasmids and *OLIG2* mRNA levels were measured using RT-qPCR. Plasmids containing individual gRNAs (S5, S6 or S7) triggered on average 36-fold (g181), 30-fold (g182) and 3-fold (g183) mRNA change, respectively. By contrast, plasmid S1 – containing all three gRNAs – activated transcription to the highest level (average of 309.8-fold change). Plasmid S2, which differed from S1 in lacking g183 achieved on average 173-fold change which is a 56% less activation compared to S1 (Fig. 3-6C). The differences between these distinct combinations are not explained by additive effects, and instead suggest synergy.

Do these synergistic effects arise because the majority of cells have consistently higher levels, or instead, are there a small number of cells that activate to a very high level? Does the presence of additional gRNAs increase the number of cells

activating the target genes, or is it the same number but with increased levels of activation at an individual cell level? Or are both these effects observed? To address these questions we used image cytometry to score OLIG2 protein in individual cells. Importantly, we have checked transfection efficiencies by staining for dCas9 positive cells and saw that all cells have similar amounts of dCas9 positive cells (Fig. 3-7 C).

With delivery of multiple gRNAs we clearly see more cells being activated within the population: S1 plasmid gives on average 52.9% OLIG2-positive cells, while S5 gives on average 35.5% OLIG2-positive cells (Fig. 3-6 D). This suggests that increased diversity of gRNAs increases the chances that sTFs can activate expression. The relative amounts of OLIG2 protein (OLIG2 immunostaining intensity) in the OLIG2-positive cells was also increased by increasing the number of gRNAs (Fig. 3-6 E). This data suggest that delivery of multiple gRNAs results in both activation of a larger proportion of cells, but also, that these will then have greater chance of higher levels of expression.

We have also explored whether there was a correlation between absolute levels of dCas9 in individual cells and the level of OLIG2. Transfected HEK293T cells were co-stained for the HA tag (HA-dCas9-VP160) and OLIG2. Interestingly, there was only weak positive correlation between level of dCas9 and level of OLIG2 (Fig. 3-7 B; weak positive correlation; Spearman $r = 0.5674$). Evidently, some cells which have high dCas9-VP160 levels do not activate *OLIG2* transcription, while some cells which have very low dCas9 levels activate *OLIG2* transcription to high levels (Fig. 3-7 A-B). This suggests that dCas9-VP160 protein levels are not a limiting factor for efficient gene activation. Rather having a diversity of gRNAs enables higher proportions of cells to activate the target gene and at higher levels.

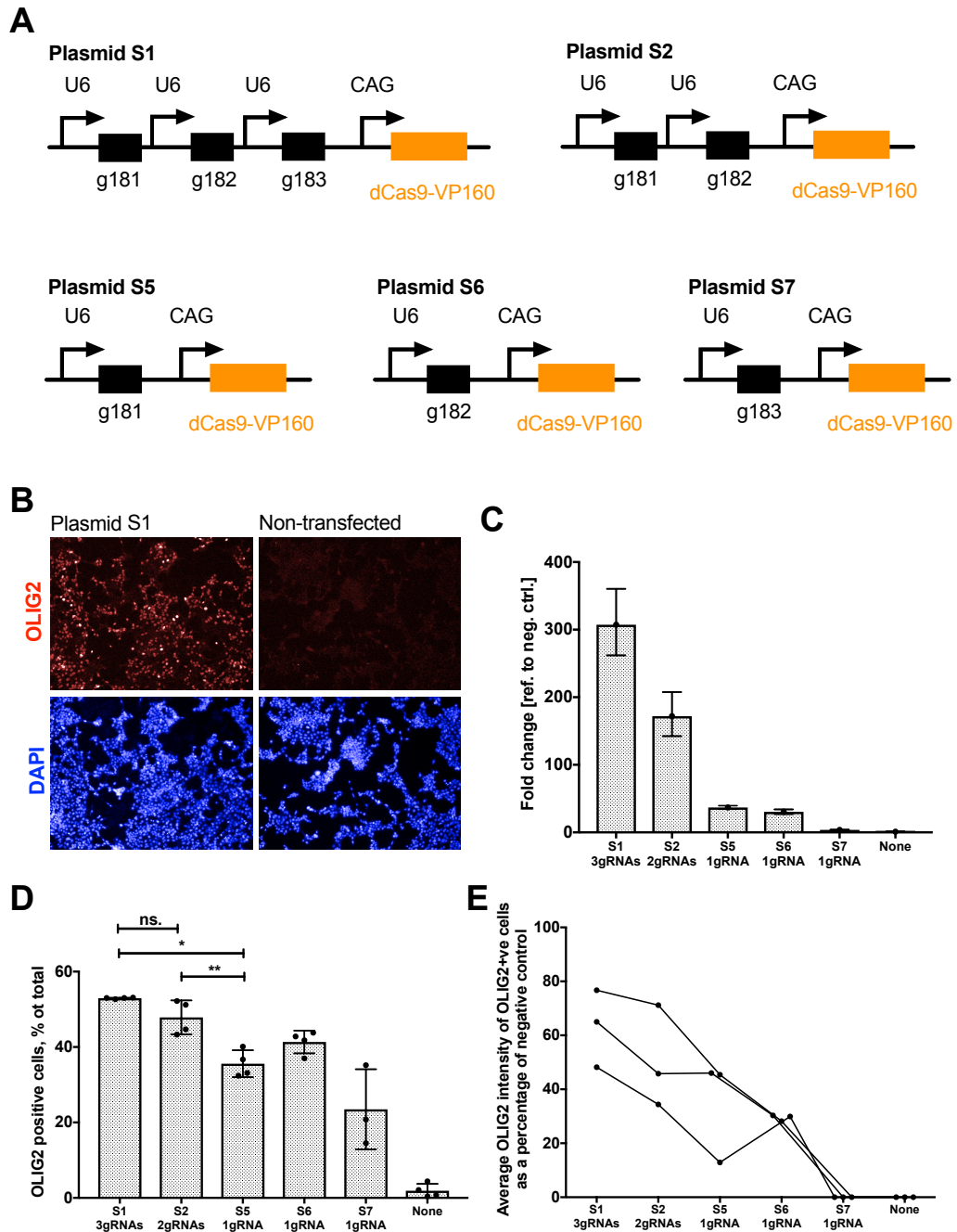


Figure 3-6 | *OLIG2* activation in HEK293T using multiple gRNAs.

(A) Graphical representation of plasmids constructed for testing the effects of co-delivery of multiple gRNAs. (B) Immunocytochemistry of OLIG2 in HEK293T cells 3 days after transfection with plasmid S1. (C) qPCR of HEK293T cells 3 days after transfection with one of the plasmids, n=1. (D) Quantification of OLIG2 positive HEK293T cells 3 days after transfection with one of the plasmids, n=4 (two-tail t-test; p-values: * < 0.001, ** = 0.005). (E) Increases in the average OLIG2 intensity in OLIG2-positive cells after transfection with different plasmids. Data shown as a percentage change in comparison to negative control, n=3; each dot represents biological replicate.

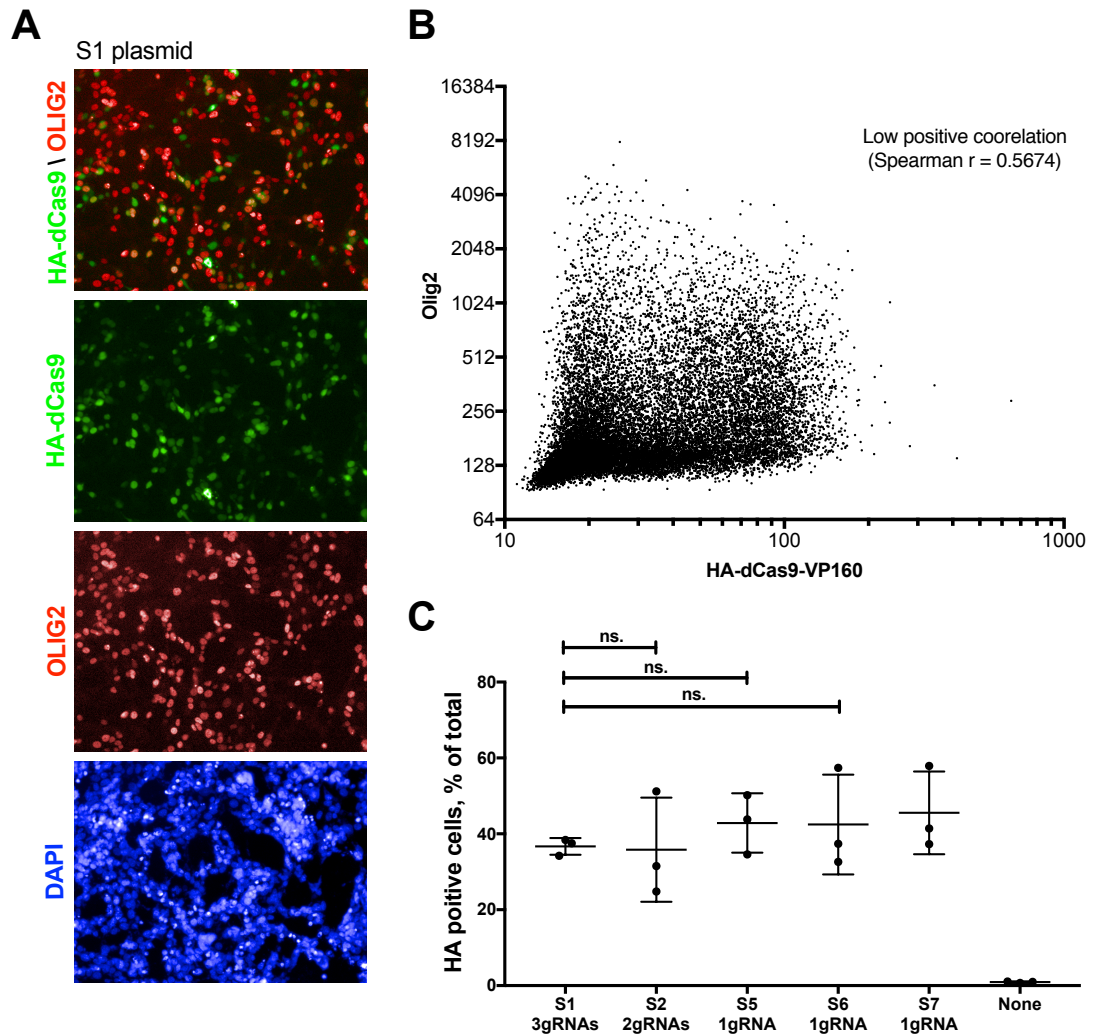


Figure 3-7 | HA-dCas9-VP160 levels in transfected HEK293T.

(A) Immunocytochemistry of OLIG2 and HA-dCas9-VP160 in HEK293T cells 3 days after transfection with plasmid S1. (B) Image cytometry quantifying protein levels for dCas9 and OLIG2 in individual cells after transfection with the S1 plasmid (three gRNAs). Each dot represents a cell. Spearman correlation analysis was performed. (C) Quantification of HA-dCas9-VP160 positive HEK293T cells after transfection with one of the plasmids, $n=3$ (S1 and S2 two-tail t-test p -value = 0.919; S1 and S5 two-tail t-test p -value = 0.260; S1 and S6 two-tail t-test p -value = 0.495).

3.2.3 Quantification of absolute mRNA levels within single cells using RNA-FISH

To further investigate how dCas9 and target gene levels correlate, we have used an RNA-FISH technique. This technique enables visualization and counting of individual mRNA molecules in each cell by the process of RNA hybridization (Fig. 3-8 A). 30 to 40 unique fluorescent probes are designed to target specific mRNA sequence. Cells are then imaged at high magnification and Z-stack images are collected. After image deconvolution, individual fluorescent dots, which represent single mRNA molecules, can be quantified to score absolute levels of mRNA (Raj & Tyagi 2010).

To first establish the procedure, we have used MEFs and probes against *Gapdh* (Fig. 3-8). Using this method, we could calculate that on average, the MEF cell contains ~2000 *Gapdh* mRNA molecules/cell (Fig. 3-8 E)

3.2.4 dCas9 and *Sox10* mRNAs molecules do not always co-localize in single MEF cells

The immunocytochemistry for OLIG2 protein levels following sTF activation in HEK293T suggested that HA-dCas9-VP160 and OLIG2 protein levels correlate only weakly (Fig. 3-7). However, this disparity might be explained by posttranscriptional regulatory differences between the two proteins. We therefore wanted to see if the same trend was observed at the mRNA level. For this purpose, we transfected MEF cells with dCas9-VP160 and 10 gRNAs targeting *Sox10* gene (g1-g10). Three days after transfection cells the processed for imaging using RNA-FISH protocol (see Methods chapter for details) (Fig. 3-9). We could identify cells which

were positive for both *Sox10* and dCas9 mRNA (Fig. 3-9 A). However, consistent with the results we saw for OLIG2 immunocytochemistry in HEK293T, we also noted little correlation between absolute levels of *Sox10* and dCas9-VP160 mRNA levels in MEFs (Fig. 3-9 B). Cells which had low levels of dCas9 mRNA sometimes had high levels of *Sox10* mRNA (cell #3-5) and vice versa: low levels of dCas9 and low levels of *Sox10* (cell #8-10), low levels of dCas9 and no *Sox10* (cell #11-12) and high levels of both mRNAs (cell #1) (Fig. 3-9 B).

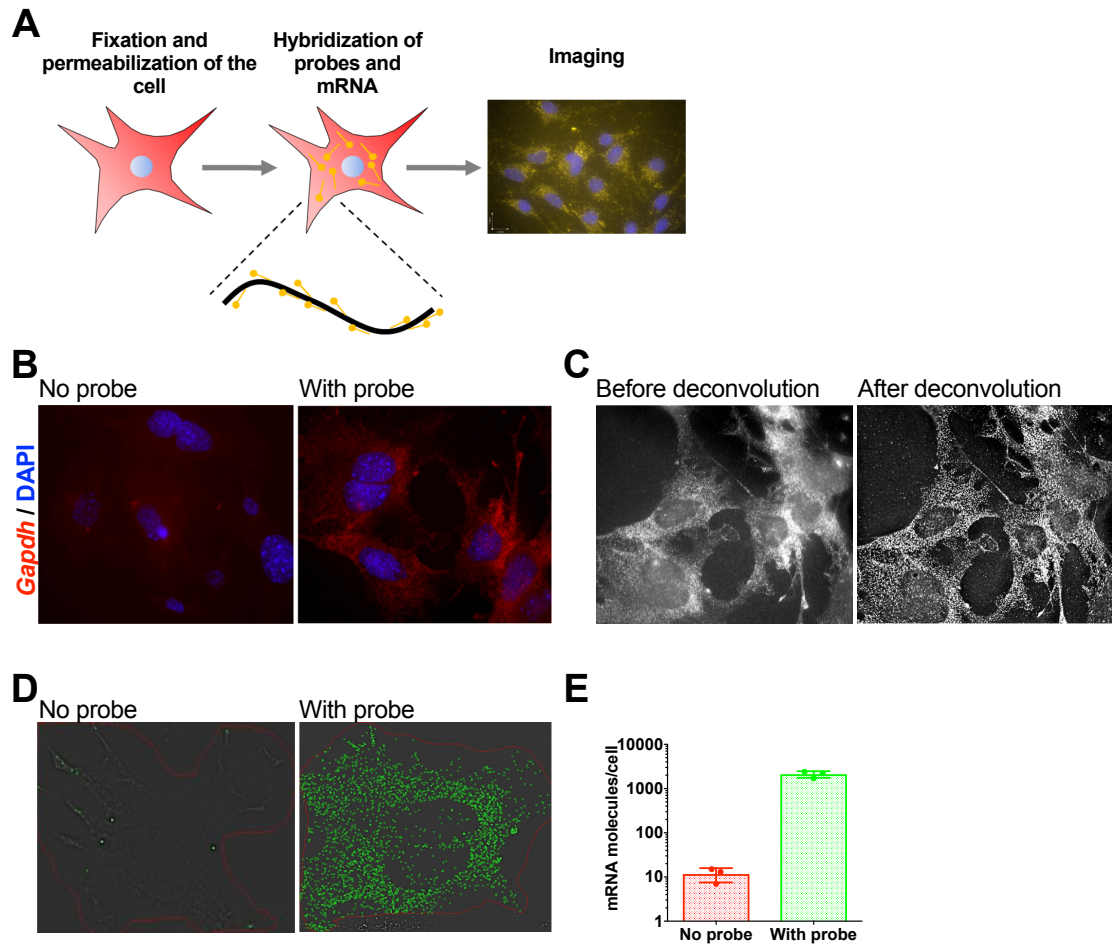


Figure 3-8 | Workflow and optimization of RNA-FISH using *Gapdh* probes in MEFs.

(A) Graphical representation of how RNA-FISH works. Cells of interest are fixed and permeabilized. Probes against specific mRNA transcript are added to cells to allow hybridization. Cells then are imaged at 60X magnification and images are processed for analysis. (B) Images obtained with Nikon (oil, 60x) demonstrating samples with and without probes against *Gapdh* mRNA transcript in MEFs. (C) Images demonstrating how the process of deconvolution enhances image resolution and combines Z-stacked image layers. (D) Example image of scoring mRNA molecules in samples treated and non-treated with probes. (E) Quantification of *Gapdh* mRNA molecules per MEF cell, n=3 cells.

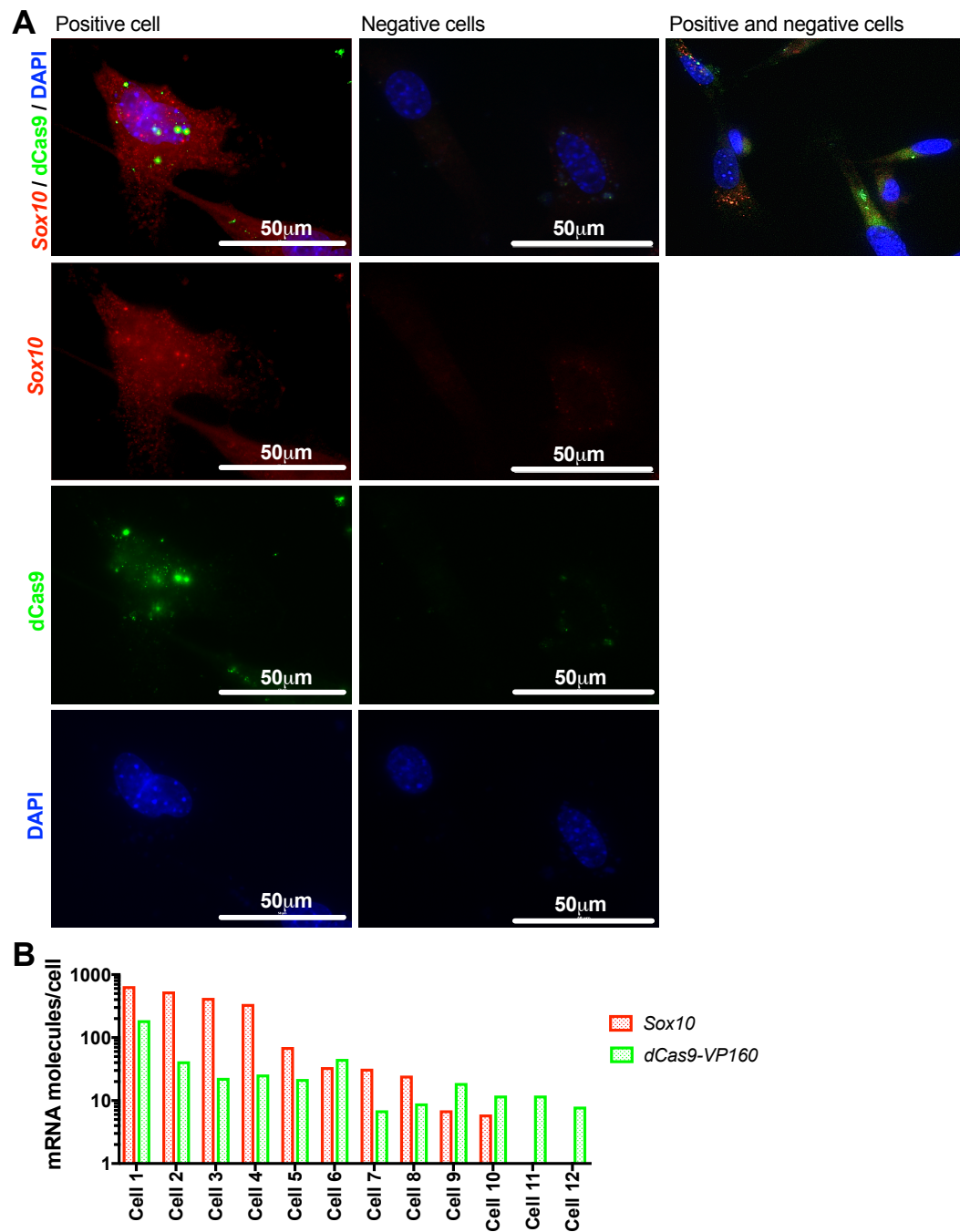


Figure 3-9 | Co-RNA-FISH for *Sox10* and dCas9 in MEFs after activation of *Sox10* with sTFs.

(A-B) Images of MEFs 3 days after transfection with dCas9-VP160 and 10 gRNAs targeting *Sox10* proximal promoter (g1-g10). MEFs were stained for probes against *Sox10* (red) and dCas9-VP160 (green) mRNA. (B) Quantification of absolute *Sox10* and dCas9-VP160 mRNA numbers in 12 cells.

3.3 *Olig2* activation in differentiating neural stem cells

3.3.1 Characterisation of a PDGFR α -GFP (PG1.1) cell line

It was evident from initial experiments that increasing expression levels of *Olig2* beyond levels of endogenous transcription is not possible in NSCs. A negative correlation between activation potential and basal expression has been reported by other groups as well (S. Lin et al. 2015). However, it is known that *Olig2* levels drop during NSC differentiation (Fig. 3-11 D). We wondered, therefore, whether activation of *Olig2* during NSC differentiation might be effective in blocking the normal downregulation that occurs, and biasing the differentiation toward OPC differentiation. A recently established PDGFR α -GFP reporter cell line (PG1.1) was used to assay OPC commitment as PDGFR α is a known marker of OPCs (Fig. 3-10).

PG1.1 reporter NSCs were derived by previous PhD student Dr. Bartłomiej Baranowski by using foetal forebrain NSCs from an established mouse strain (Hamilton et al. 2003). This mouse strain was originally constructed by inserting H2B histone protein tagged with GFP (H2B-GFP) under the control of PDGFR α (Fig. 3-10 A) (Hamilton et al. 2003). As a result, after PG1.1 cells are forced to differentiate to OL lineage, some cell nucleus becomes GFP positive (Fig. 3-10 B-D). This can be scored by live imaging (Fig. 3-10 C), flow cytometry (Fig. 3-10 D) and immunocytochemistry (Fig. 3-11 B). As PDGFR α is not expressed in undifferentiated NSCs and is absent from astrocyte and neuronal cells, this NSC cell line is a useful reporter for OPC lineage. This was confirmed, as OPC markers such as *Olig2* and *Sox10*, are highly enriched in the GFP positive cells after four days of differentiation, sorting and qPCR (Fig. 3-10 E.)

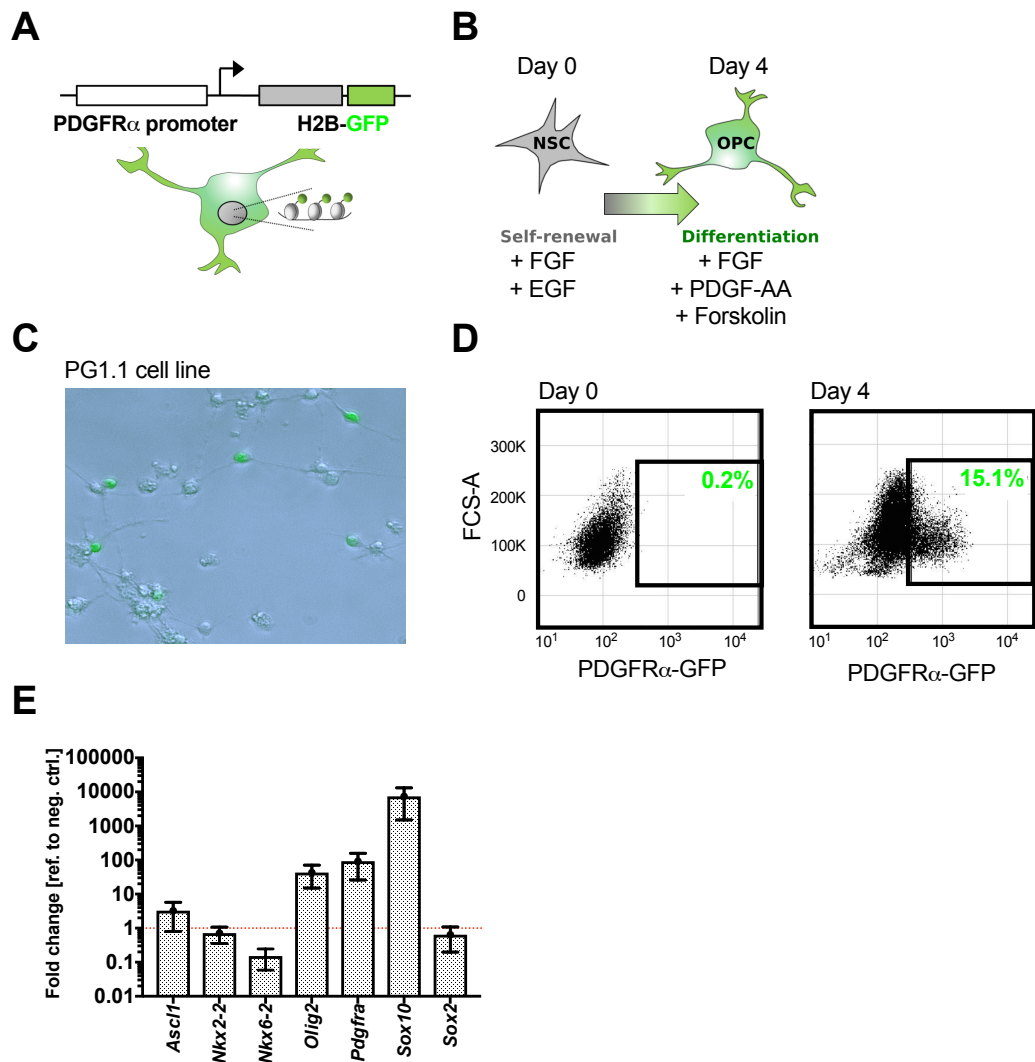


Figure 3-10 | Introduction to PDGFR α -GFP (PG1.1) cell line.

(A) Schematic of the PG1.1 reporter mouse NS cell line. PG1.1 cells were originally derived by Dr. Bartłomiej Baranowski (Pollard lab, unpublished). (B) Schematic of PG1.1 differentiation and activation of the GFP reporter. Under self-renewal conditions (presence of both FGF-2 and EGF) cells lack fluorescence; however, upon differentiation (removal of EGF and addition of FGF-2 plus PDGF-AA and Forskolin; 4 days) cells upregulated PDGFR α and become GFP positive. (C) Live phase contrast and fluorescence overlay of PG1.1 cells 4 days after differentiation. (D) Flow cytometry analysis of differentiation (Day 4) of PG1.1 cell line (day 0 is control, left panel). (E) qPCR data showing mRNA levels for various OPC markers in GFP positive cells after differentiation (cells were sorted for GFP 4 days after differentiation), n=1.

3.3.2 *Olig2* activation during PG1.1 differentiation to OPCs or astrocytes

We first tested that PG1.1 cells can be differentiated reliably in the absence of sTFs. As anticipated, during differentiation a proportion of cells extinguished OLIG2 expression, as evident by immunostaining (Fig. 3-11 D). Quantification of levels of OLIG2 using image cytometry was then tested. Prior to differentiation (black dots) all NS cells are OLIG2 positive and PDGFR α -GFP negative. Following differentiation (green dots), the population splits into at least two compartments including the double positive OLIG2/PDGFR α -GFP (Fig. 3-11 B).

We next tested if delivery of the sTF to activate *Olig2* can work during differentiation (i.e. differentiation was triggered immediately after transfection). We hypothesized that it might be possible to activate *Olig2* in cells which lose OLIG2 levels during differentiation. To test if *Olig2* can be activated in differentiating NSC cells, the PG1.1 reporter cell line was transfected with dCas9-VP160 and gRNAs targeting *Olig2* (g11-g20). These were compared to a control cells which were transfected with only dCas9-VP160. Four days after starting differentiation, these cells were stained for OLIG2 and GFP and analysed by image cytometry (Fig. 3-11 B-C).

After repeating and quantifying 4 independent differentiations/transfections, we could not see any statistically significant increase in the proportion of OLIG2 positive cells with the presence of sTF (Fig. 3-11 E-F). However, in three out of four experiments there was a trend to have an increase in the proportion of OLIG2 positive cells within the experiment (Fig.3-11 E, lines represent related experiments).

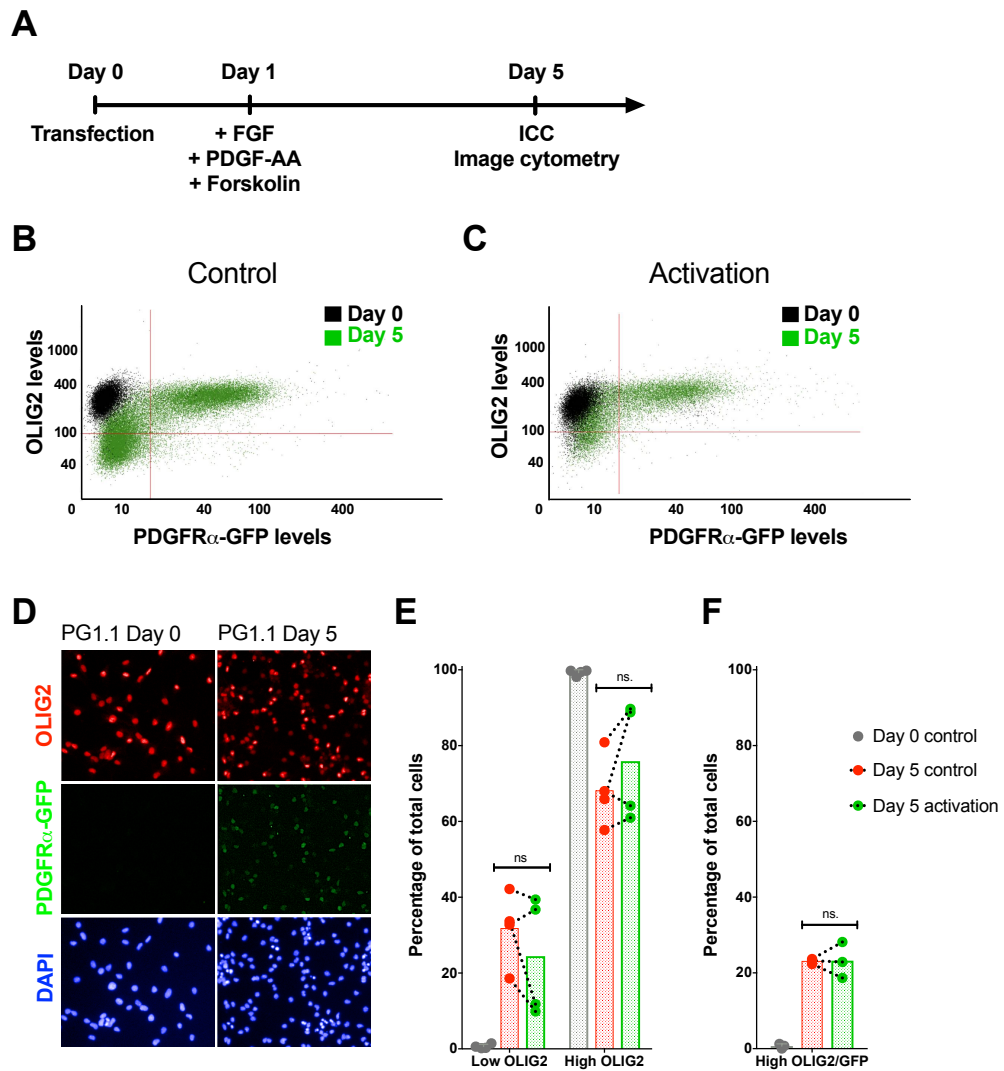


Figure 3-11 | *Olig2* activation during PG1.1 differentiation to OPCs.

(A) Outline of experimental workflow. Cells were transfected with dCas9-VP160 and 10 gRNAs for *Olig2* (g11-g20) or dCas9-VP160 alone (control). The next day cells were induced to differentiate for 4 days and then ICC was performed for OLIG2 and GFP (for PDGFR α -GFP) levels; data was quantified using image cytometry. (B) Results of image cytometry: each dot represents a single cell, visualizing levels of OLIG2 and PDGFR α -GFP before (black dots) and after (green dots) differentiation. Cells transfected with dCas9-VP160 alone (“Control”). (C) Cells transfected with dCas9-VP160 and 10 gRNAs for *Olig2* activation (“Activation”). (D) Immunocytochemistry staining for OLIG2 and PDGFR α -GFP in the PG1.1 cell line before (Day 0) and after differentiation (Day 5). (E) Quantification of image cytometry data demonstrating the percentage of OLIG2 low and OLIG2 high cells in each condition. Lines link data from same experiments, n=4. (F) Quantification of image cytometry data demonstrating the percentage of OLIG2/GFP high cells in each condition. Lines link data from same experiments, n=3.

An alternative differentiation regime for NS cells is to expose them to BMP4 at low density. This triggers a uniform differentiation to astrocytes; >95% of cells express GFAP, exit the cell cycle and *Olig2* expression is extinguished in all cells (Fig. 3-12 C). We tested if activating *Olig2* with the sTFs can override the astrocyte promoting effects of BMP4. Cells were transfected either with dCas9-VP160 alone or with a pool of 10 gRNAs (g11-g20) and dCas9-VP160 and induced to differentiate next day by BMP4 exposure (Fig. 3-12 A). At Day 4 (three days of differentiation), cells were analysed by qPCR and immunocytochemistry. We found no difference in *Olig2* transcription following delivery of sTFs at mRNA (Fig. 3-12 B) or protein level (Fig. 3-12 E). Altogether, these results suggest that *Olig2* sTFs are not potent enough to override the differentiation programs operating during NS cell differentiation.

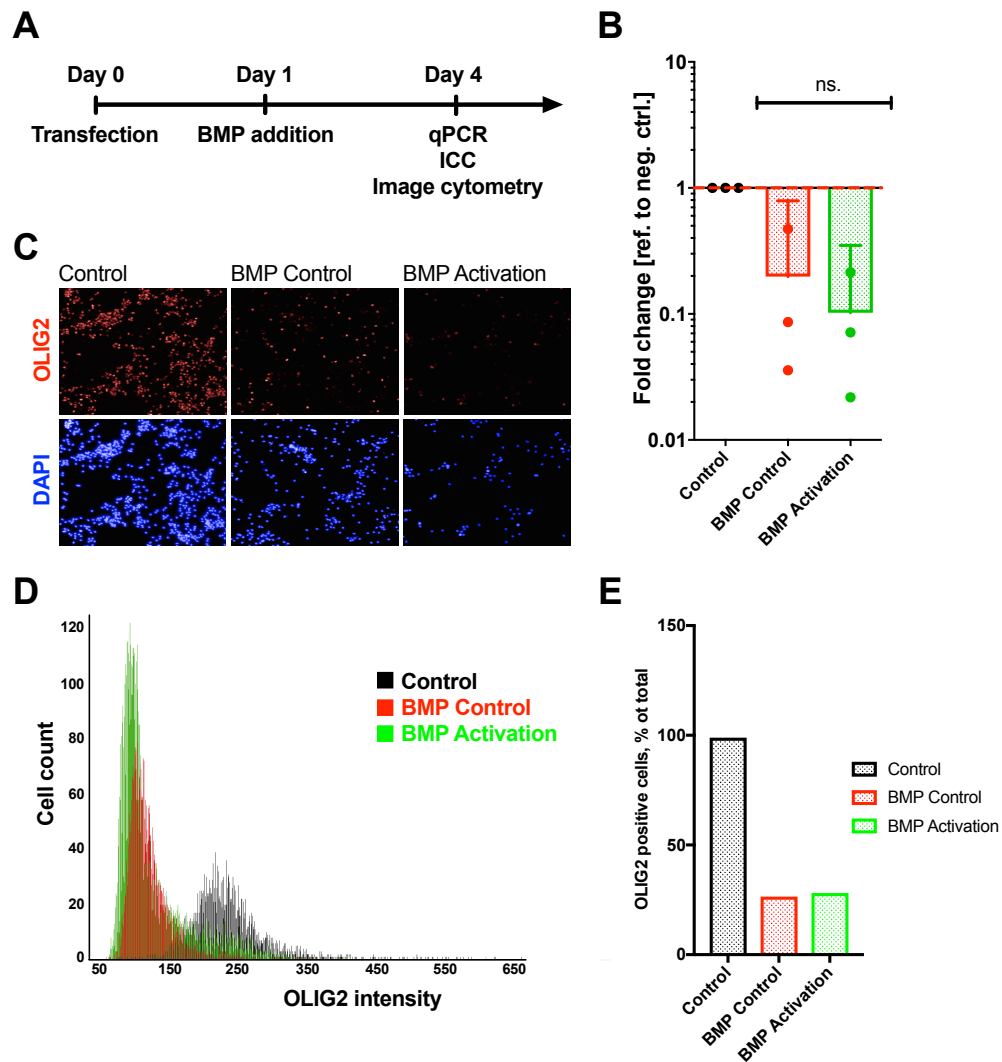


Figure 3-12 | *Olig2* activation during PG1.1 differentiation to astrocytes.

(A) Outline of experimental workflow. PG1.1 mouse NS cells were transfected with dCas9-VP160 (“Control”) alone or together with 10 gRNAs for *Olig2* (g11-g20). The following day cells were induced to differentiate for three days in presence of BMP4. Cells are imaged for OLIG2 and parallel wells were used for mRNA analysis by qPCR. (B) qPCR results 4 days after transfection with either dCas9-VP160 alone (“BMP Control”) or dCas9-VP160 and 10 gRNAs for *Olig2* (“BMP activation”) and 3 day treatment with BMP4 or no BMP4 (“Control”), n=3. (C) Immunocytochemistry for OLIG2 in BMP4-untreated non-transfected (“Control”), BMP-treated transfected with dCas9-VP160 alone (“BMP control”) and BMP-treated transfected cells with dCas9-VP160 and *Olig2* gRNAs (“BMP Activation”). (D) Image cytometry graph showing quantification of OLIG2 immunocytochemistry of treated and non-treated cells, (E) Quantification of (d), n=1.

3.4 Investigating potential of different transcription activating domains (TADs) to overcome biological restrictions for activation

3.4.1 Screening of new transcription activating domains (TADs)

The results in chapter 3.3 suggested that sTFs can work well to activate some target genes, but have cell type dependent effects. For example, *Nkx6-2* could be easily activated in MEFs, but not at all in NS cells. Moreover, even when using functional gRNAs we did not observe ‘overexpression’ over physiological levels (*Olig2* in NSC). One explanation for each of these findings could be that the VP160 effector is not strong enough to achieve transcriptional activation in some situations. Alternative transcriptional activation domains (TADs) are often used and suggested to give stronger activation e.g. p300 and VPR (Chavez et al. 2016). Thus, we next explored whether novel and previously reported TADs might work well in NS cells and MEFs.

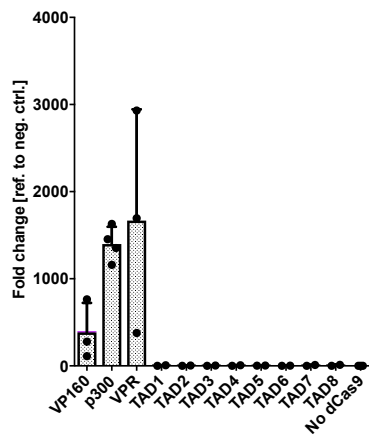
Firstly, we screened eight novel TADs for their ability to activate *Sox10* in PG1.1 NSC line, using the gRNAs already validated with VP160. TADs were designed and constructed by a postdoc in the Pollard lab (Dr. Andrea Martella). These contained various versions and combinations of previously reported domains that have a role in transcription activation (Fig. 3-13 A). To test their potential, we transfected PG1.1 NSCs with a pool of 10 gRNAs targeting *Sox10* (g1-g10), and used VP160 as a normalisation positive control. From the qPCR data, it is evident that none of the novel designed TADs were able to activate *Sox10*. However, p300 and VPR effectors outperformed VP160 as expected and showed significantly higher level of activation: more than 3 times over VP160 (Fig. 3-13 B).

Next, we were interested to see if stronger effector (p300 and VPR) would be able to overcome biological restrictions in NSC and activate *Nkx2-2* and *Nkx6-2*. Both of these genes could be activated in MEFs (Fig. 3-13 D and Fig. 3-3), which means that gRNAs are functionally capable to activate transcription. However, *Nkx6-2* could not previously be activated in NSCs with VP160. PG1.1 NSC were transfected with either a pool of gRNAs targeting *Sox10* promoter, *Nkx2-2* promoter, *Nkx2-2* enhancer or *Nkx6-2* promoter and either VP160 (green), p300 (red) or VPR (blue). Neither *Nkx2-2* or *Nkx6-2* could be activated in NSC. Moreover, even activation of *Nkx2-2* in MEFs is not significantly enhanced by using stronger effector (Fig. 3-13 D). Thus, we confirmed that alternative activators can work more effectively than VP160 in some cases, however, effector strength does not seem to be the explanation for the failure to activate some other genes. Instead, some difference in promoter use or epigenetic status of these loci between cell types could explain the lack of activation.

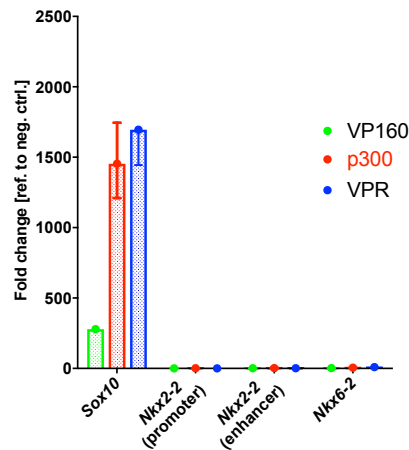
A

TAD	Encoded protein	Amino acid sequence	Reference
TAD1	Modified GCN4	KTVLPIPELDDAWWWLFWFS	PNAS (2014) 13:E3506–E3513
TAD2	KIX domain interacting with p300/CBP	SWAVYELLF	Nat. Biotechnol. (2000) 18 :1080–1085
TAD3	Modified domain from p53	LTEELIELWFTE	J. Am. Chem. Soc., 2012, 134, 1715–1723.
TAD4	Chromatin unfolding peptide CUP1.1	DELQPASIDP	Molecular and cellular biology, Feb. 2005, p. 958–968
TAD5	TAD3 combined with CUP1.1		
TAD6	4 versions of TAD1		
TAD7	4 versions of TAD2		
TAD8	4 versions of TAD3		

B



C



D

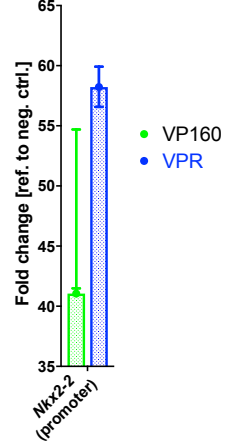


Figure 3-13 | Screening of new transcription activating domains (TADs).

(A) Table summarising the types of novel TADs that were explored. TADs were constructed by Dr. Andrea Martella (Pollard lab). (B) qPCR data showing *Sox10* transcription levels in the PG1.1 cell line 3 days after transfection with 10 gRNA (g1-g10) and dCas9 containing one of the specified TADs, n=2. (C) qPCR data showing transcription levels of each target gene in PG1.1 cell line 3 days after transfection with either dCas9-VP160 (green), dCas9-p300 (red) or dCas9-VPR (blue) and pool of 10 gRNAs targeting named genes, n=1. (D) qPCR data showing transcription levels of *Nkx2-2* genes in MEFs 3 days after transfection with either dCas9-VP160 (green) or dCas9-VPR (blue) and pool of 10 gRNAs targeting *Nkx2-2*, n=1.

3.5 Mouse gene repression in NSC using dCas9-KRAB

3.5.1 gRNA placement for gene repression

In addition to transcriptional activation, transcriptional repression using sTFs has also been reported by a number of groups. These are created by fusing repressor domains such as KRAB to the dCas9 (Gilbert et al. 2014). We were interested in testing whether combining activation of some targets with repression of others would be possible, as this might improve the efficiency of programming or reprogramming cell fate. For this purpose, I selected the *Sox9* gene as one of the candidates for testing repression, due to its reported role in repressing *Sox10*, a known master regulator of OPC cell fate (C. D. Pozniak et al. 2010b). Such repression, if successful, might help drive terminal differentiation of OPCs.

Genome-wide libraries have been used to define the design rules for effective repression using sTFs with KRAB domains (Gilbert et al. 2014). gRNAs were found to be effective when targeted to -50 to 200bp around the TSS. Six gRNAs, spanning this region for *Sox9* were cloned into expression vectors using the method described in Figure 3-2 (Fig. 1-14).

3.5.2 *Sox9* repression in SOX9-V5 NS cell line

To test if repression of *Sox9* is possible, we made use of a SOX9-V5 NS cell line (constructed by Dr. Pooran Dewari). After optimizing the transfection efficiency for this new cell line (Fig. 3-15 A), cells were either transfected with dCas9-KRAB and 6 gRNAs targeting the proximal *Sox9* promoter, or with dCas9-KRAB alone. 3 days after transfection, cells were analyzed by V5 immunocytochemistry and qPCR (Fig. 3-15).

After quantifying SOX9-V5 low and high cells, as defined by the specific intensity threshold (Fig. 3-15 C), it was evident that after transfection there was indeed an increase in the V5 low/negative population in cells transfected with gRNAs compared to cells which only got dCas9-KRAB (Fig. 3-15 D). This data suggested that dCas9-KRAB repress SOX9 in approximately 10% of the cells.

However, this could not be confirmed by qPCR 3 days after transfection, as no difference between two conditions could be observed (Fig. 3-15 E). It is not clear why this is the case, but similar observation when dealing with repression was also made by previous PhD student in the lab as well. The repressors were not explored further in the programming and reprogramming experiments in chapter 5.

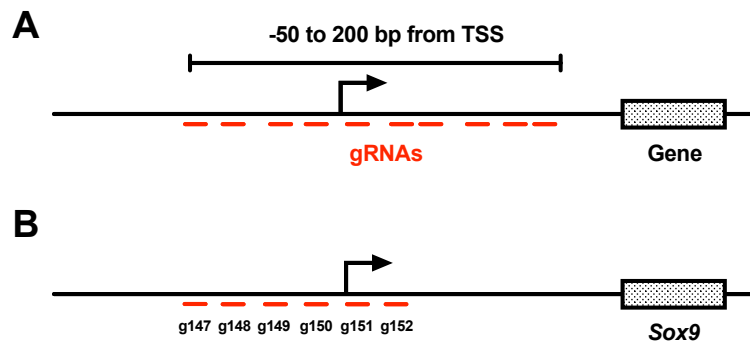


Figure 3-14 | gRNA placement for *Sox9* gene repression.

(A) Strategy for a design of gRNAs for transcriptional repression. (B) gRNA positioning along *Sox9* proximal promoter and transcription start site (TSS) for transcription repression.

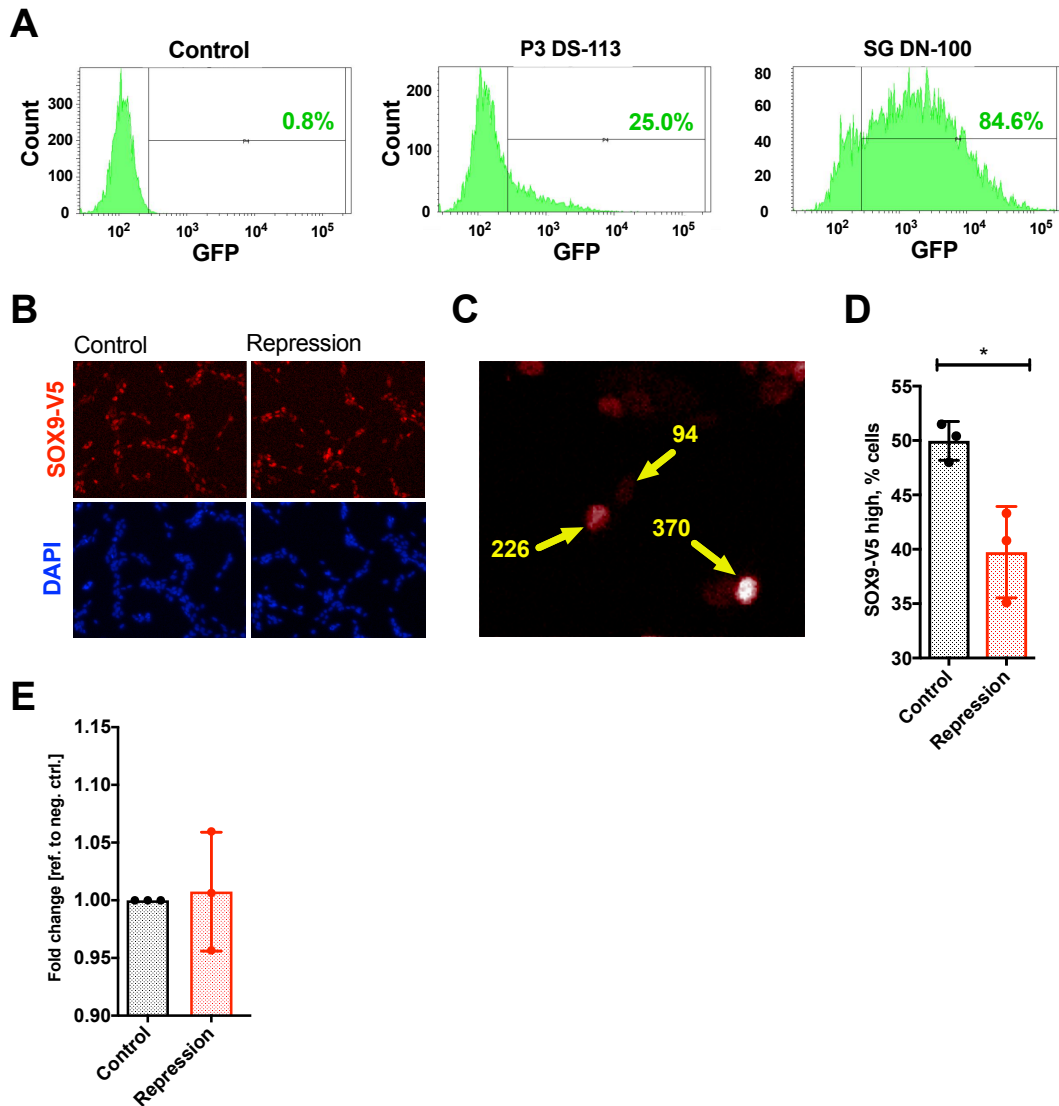


Figure 3-15 | *Sox9* repression in SOX9-V5 cell line.

(A) Nucleofection protocol optimization for the SOX9-V5 cell line using a GFP reporter plasmid. The SOX9-V5 cell line was constructed and validated by Dr. Pooran Dewari (Pollard lab). (B) Immunocytochemistry for V5 in SOX9-V5 cells after transfection with dCas9-KRAB alone (“Control”) or dCas9-KRAB and 6 gRNAs (g147-g152) targeting *Sox9* transcription start site (“Repression”). (C) Image of stained V5 in SOX9-V5 cells to demonstrate gating strategy based on SOX9 intensity. Anything below intensity threshold of 95 is considered as SOX9-V5 low/negative and vice versa. (D) Quantification of image cytometry data showing percentage of SOX9-V5 high cells with (“Repression”) and without (“Control”) repression, n=3 (two-tail t-test, p-value = 0.017). (E) qPCR data of *Sox9* mRNA levels 3 days after cells were transfected with either dCas9-KRAB alone (“Control”) or with dCas9-KRAB and 6 gRNAs targeting *Sox9* TSS (“Repression”), n = 3.

3.6 Concluding remarks

In this Chapter I have explored whether sTFs can be used in MEFs and NS cells to activate genes of interest. Several important findings and conclusion have been made. Firstly, we show that some gRNAs can function individually in successfully activating gene transcription and that large difference in potential to activate transcription exists even between very proximal gRNAs. Moreover, pools of gRNAs give higher transcription activation compared to single functional gRNAs. This is important finding because level of gene transcription activation might be important to successfully modulate biological functions. Interestingly, a combination of gRNAs which individually are minimally functional can still function well in a “pool”, suggesting that synergy between gRNAs exist. After we looked closer into synergy between gRNAs, we have surprisingly found that combining functional gRNAs together increase a number of cells which are activated and also level of activation within those cells.. Lastly, we demonstrate evidence that suggest that levels of dCas9-VP160 are not limiting factor for successful transcription activation. After single cell image cytometry analysis, we report that weak positive correlation exists between target gene levels and dCas9-VP160 levels, suggesting that other factors (like gRNA concentration and variation) play more important role in gene activation.

Furthermore, we have shown that different effectors can have significantly different activation strengths. As observed by *Sox10* transcription activation in NSC, p300 and VPR were capable of significantly higher level of transcription activation compared to VP160. That being said, this is not always a case as *Nkx2-2* activation with VP160 and VPR in MEFs has shown. Evidently, we demonstrate that even strongest effector (VPR) might not be able to significantly outperform weaker VP160

effector in some contexts. This is important observation as it shows that other potentially limiting factors need to be explored instead (such as alternative promoters).

Thirdly, in addition to technical limitations, biological restrains seem to play a significant role in gene activation using sTFs. Firstly, genes cannot be over-activated if it is already transcribed, as we have seen with *Olig2* in NSC. Importantly, we looked into the prospect of such gene transcription activation in situations in which its' transcription levels are lowered – e.i. during differentiation – which has not been investigated before. Surprisingly, even if levels of transcription are lowered in these cases (via differentiation), we could not find statistically significant differences (on average) between control and treated cells. That being said, when we looked into individual experiments, we saw that in 3 out of 4 experiments we detected more OLIG2 positive cells after treatment. Secondly, additional biological constrains exist that makes gRNAs which are functional in one cell line, non-functional in another cell line as we have observed with *Nkx6-2* activation in NSCs and MEFs. These biological constrains cannot be overcome by using stronger effectors which might indicate that dCas9/gRNA binding to DNA could be a key limiting factor.

4 Multiplex delivery of sTFs: a new toolkit for plasmid assembly

4.1 Design and validation of an all-in one plasmid for multiplex transcriptional activation

In chapter 3 our observations suggested it would be beneficial to deliver multiple gRNAs simultaneously to achieve efficient transcriptional activation. Such multiplexing may also be required for multiple distinct targets genes to be activated concurrently. This would typically require co-transfection of tens of plasmids simultaneously and would likely result in a heterogeneous population of cells with variable amount and proportions of each plasmid (Fig. 4-1A).

This problem was highlighted to us when we directly tested co-transfection of two separate plasmids constitutively expressing either GFP and mCherry using the Amaxa Nucleofection. Analysis using flow cytometry unexpectedly revealed that the majority of cells only expressed one or the other plasmid; a smaller fraction than anticipated were co-expressing (Fig. 4-1B). Thus, it is easy to predict that by increasing a number of plasmids that is being transfected, the ratio of cells expressing all plasmids would most likely decrease even further – limiting the success of multiplexing using transient transfection.

4.1.1 Designing and validating a novel extensible Mammalian Modular expression plasmid Assembly (EMMA) system

To tackle the multiplexing problem and make screening of different effectors and gRNAs easier, we developed a novel system for plasmid assembly which we named “Extensible Mammalian Modular Assembly” or EMMA (Martella et al. 2017).

The overall EMMA approach was developed by Andrea Martella, a postdoc in the Pollard/Cai laboratory. I worked closely with Andrea to exploit EMMA and to validate many of the newly ‘domesticated’ core expression vector parts (Martella et al. 2017). I performed the experiments to assemble and test sTFs using EMMA.

EMMA employs Golden Gate (GG) restriction enzyme/ligation reactions to enable large number of compatible vector parts to be assembled in a single step reaction. The GG reaction is based on the use of type IIS restriction endonucleases, ligase and multiple rounds of digestion and ligation.

A large library of parts from existing vectors were ‘domesticated’, i.e. synthesized to remove type II sites and create the appropriate overhangs and linkers. These parts included: various promoters (CAG, EF1, U6 etc.), fluorescent proteins (mRuby, NeoGreen etc.), localization signals, tags etc. Each part is surrounded by a BsmBI restriction site with a unique overhang (Fig. 4-2A). This allows parts to be assembled in a specific order which is important for functionality; e.g. the promoter always sits adjacent to the coding sequence. These parts were cloned into intermediate plasmids and stored for later use.

To make a new plasmid using EMMA, the appropriate parts are selected from the library of plasmids. These are assembled with an empty backbone plasmid in a one-step GG reaction (see Methods for details). Using this system Martella et al., were able to routinely generate expression plasmids with 15-25 different parts in a single reaction. Non-assembled backbones cannot be propagated as transformed bacteria will die due to the *ccdBS* negative selection cassette being present (Fig. 4-2B). The whole process only takes two days: one day for the GG reaction (3-8 hours depending on part numbers), and one day for transformation (Fig.4-2C). Final vectors were validated by

restriction digestion and sequencing. EMMA has been designed to accommodate up to three separate transcription units, can use polycistronic elements, and is compatible with multiple delivery methods (transposase, targeting vector, transient, viral etc.). The design of this system was made with maximum flexibility in mind.

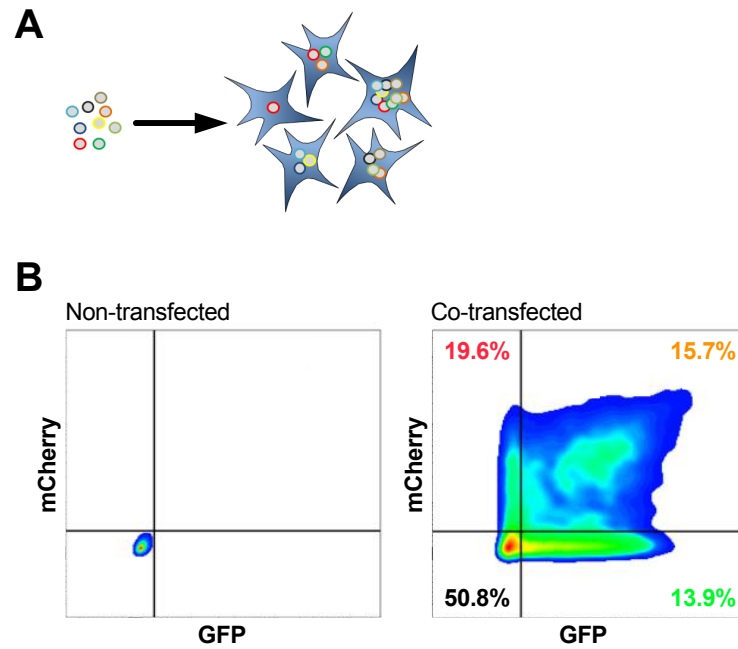


Figure 4-1 | Lack of effective co-transfection poses a hurdle for multiplex delivery of sTFs.

(**A**) Graphical representation of the problems around co-transfecting a large number of plasmids together. The resulting population is heterogeneous with cells having different proportions of each plasmid. (**B**) Flow cytometry data of PG1.1 NSCs three days after co-transfection with CAG_GFP and CAG_mCherry expression plasmids (n=2).

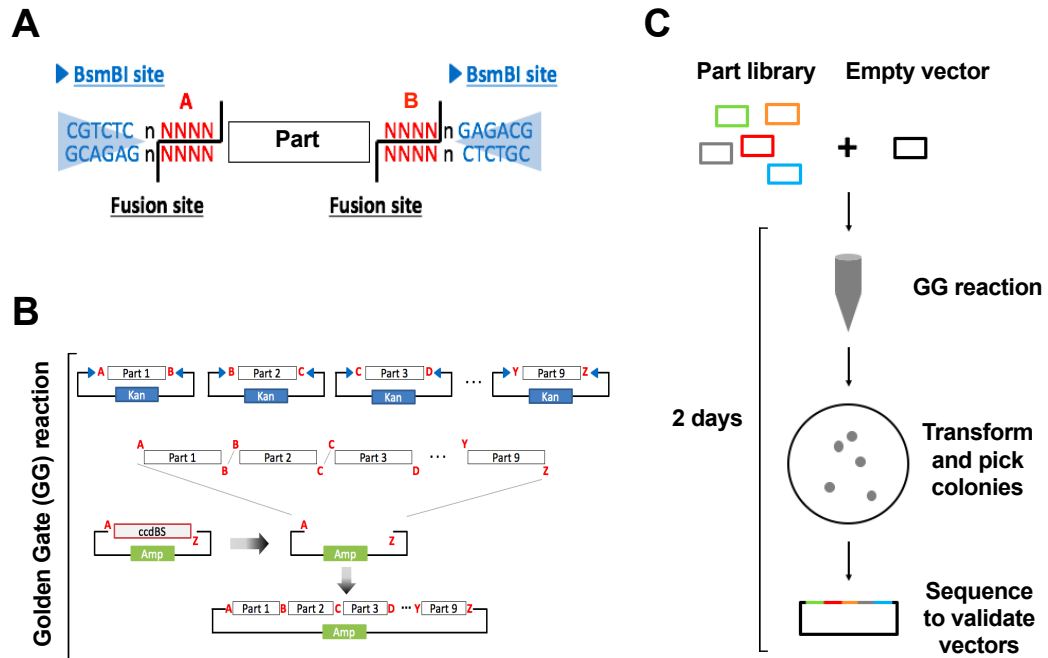


Figure 4-2 | Summary of an Extensible Mammalian Modular Assembly (EMMA) system and the process of assembly.

(A) Graphical representation of how EMMA plasmid parts of interest are surrounded by BsmBI restriction sites and unique overhangs that allow efficient assembly. (B) After parts with unique overhangs are created they are used in a Golden Gate (GG) reaction (digestion and ligation cycles). This reaction allows assembly of the desired plasmid. (C) Overview of whole process. Plasmids containing parts with specific overhangs are created (see A) which are then assembled into a full plasmid using GG reaction. Assembly products are transformed and bacteria colonies are picked. Correctly assembled plasmids are tested by restriction digestion and sequencing.

To validate that it is possible to assemble functional plasmids using the EMMA system, we first assembled a number of different expressions plasmids, including two plasmids presented in Figure 4-3 (plasmids were assembled by Dr. Andrea Martella; I performed all functional testing presented in the Figures).

The first plasmid that was built was a simple green fluorescence protein NeoGreen, tagged with membrane localization signal. This cassette was driven by a CAG promoter. HEK293T cells were transfected and live cells imaged three days later. As expected, the HEK293T cell membrane (and vacuoles) were tagged with green fluorescence protein (Fig. 4-3B).

To further validate EMMA, more elaborate plasmids were constructed. For example, a CAG promoter with two fluorescence proteins downstream – green NeoGreen and red mRuby – the former with a localization signal for mitochondria while mRuby had no localization signal. Both fluorescent proteins were driven by the same promoter but separated using the P2A translational cleavage sequence (Fig. 4-3C). As predicted, HEK293T cells transfected with such plasmid had red cytoplasm and green mitochondria (Fig. 4-3D).

I confirmed the success and utility of EMMA using many further examples in the study by Martella et al. But for simplicity I report here our efforts to use an EMMA related system for sTF expression vectors.

4.1.2 Construction of toolkit for gene regulation in multiplex

With the EMMA GG plasmid assembly system validated, we next moved on to adapt this system for construction of an ‘all-in one’ (Ai1) sTF expression plasmid which would be able to deliver multiple gRNAs from a single plasmid. Firstly, an empty Ai1 vector was made using the EMMA system (Fig. 4-4). This vector contained PiggyBac homology arms to enable stable integration, an mRFP expression cassette, dCas9-VP160 driven by the constitutive CAG promoter and an mRuby-p2A-Puro cassette driven by another CAG promoter (Fig.4-4A).

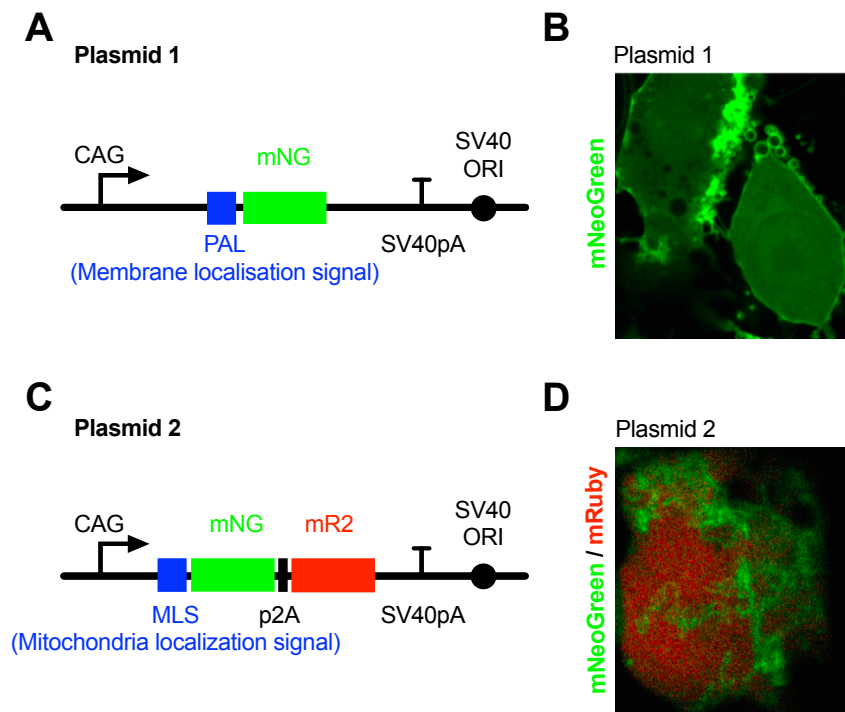


Figure 4-3 | EMMA plasmid assembly using fluorescent proteins and localization tags.

(A) Schematic representation of Plasmid 1, which was assembled with EMMA from five parts. Plasmid 1 contained constitutive promoter (CAG) NeoGreen fluorescent protein (mNG) tagged with membrane localization signal (PAL). Assembly was performed by Dr. Andrea Martelli. (B) Live imaging of HEK293T cells three days after transfection with Plasmid 1. (C) Graphical representation of Plasmid 2, which was assembled with EMMA from seven parts. Plasmid 2 contains constitutive promoter (CAG), NeoGreen fluorescent protein (mNG) tagged with mitochondria localization signal (MLS), p2A sequence and mRuby fluorescent protein (mR2). (D) Live imaging of HEK293T cells three days after transfection with Plasmid 2. These experiments were designed to confirm functionality of key components of the EMMA parts.

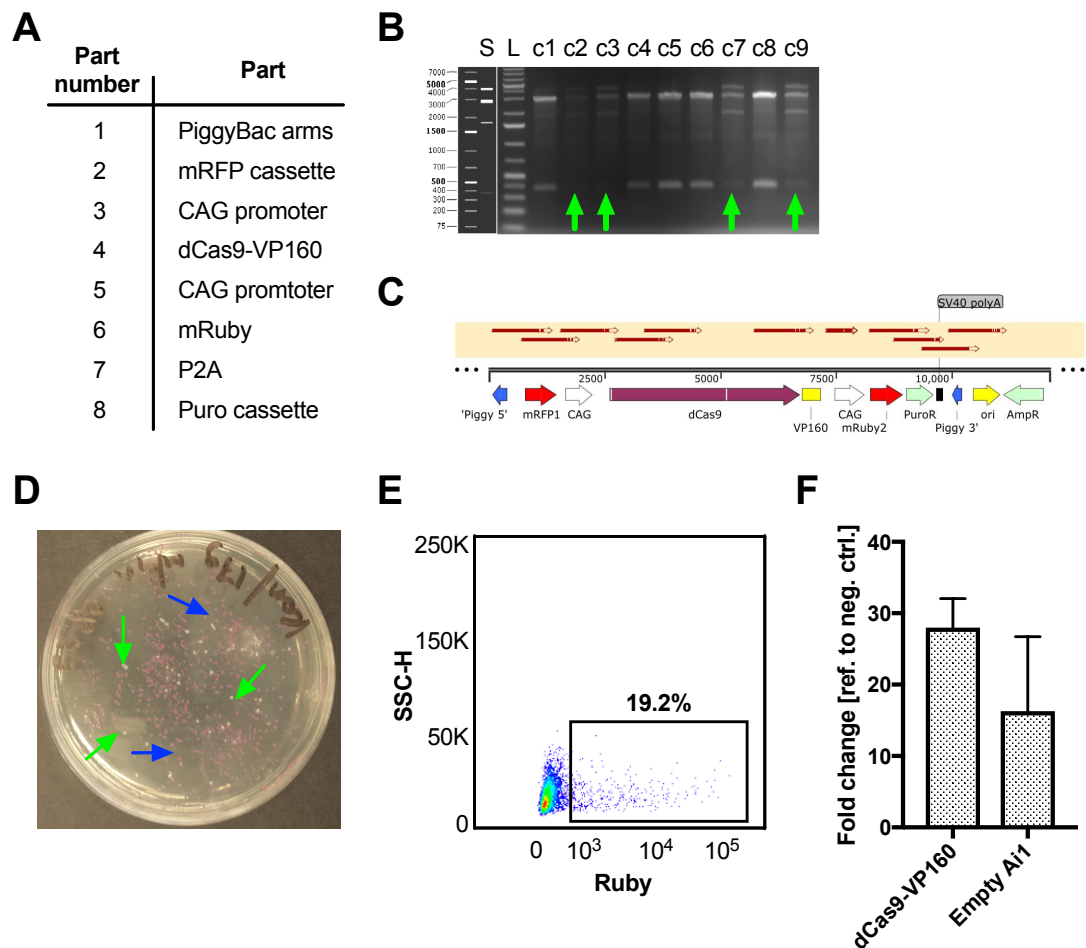


Figure 4-4 | Making and validating empty All-in one (Ai1) vector for multiplex gene activation.

(A) Table of parts that has was used to assemble the entry Ai1 plasmid capable of harboring gRNAs cloned in via replacement of the mRFP cassette. (B) Test digestion of empty Ai1 plasmid clones. Simulated agarose gel of correct clone digestion is shown on the left (S). Correct clones are indicated by green arrows. (C) Sequencing results of empty clone 2 of Ai1 plasmid showing correct part assembly. (D) A picture of agarose plate with *E.coli* transformed with Ai1 after GG reaction. Red bacteria colonies (blue arrows) indicate empty Ai1 plasmids (still contain RFP cassette between GG restriction sites), white bacteria colonies (green arrows) indicate assembled Ai1 plasmids (RFP cassette replaced by gRNAs). (E) Flow cytometry data showing intensity of mRuby reporter in PG1.1 NSC two weeks after transfection with empty Ai1 vector and PiggyBac Transposase. (F) qPCR data showing mRNA levels of *Sox10* in PG1.1 NSCs three days after transfection with a pool of 10 gRNAs targeting *Sox10* and either CAG_dCas9-VP160 or empty Ai1 plasmid, n=1.

After assembly of the empty Ai1 plasmid, multiple clones were tested by restriction digestion (Fig. 4-4B) and 4 out of 9 (giving estimated efficiency of 44%) clones were correctly assembled in this 8-part EMMA reaction. Clone 2 was sequenced to further validate correct assembly of the parts (Fig. 4-4C).

To further test if these parts are functional, a number of tests have been performed.

Firstly, the mRFP cassette was tested. Bacterial colonies transformed with an empty Ai1 had a red appearance (Fig. 4-4B). Importantly, mRFP was surrounded by BsmBI restriction sites, which allows multiple gRNA units to be added to this site using Golden Gate reaction, as long as these units have correct overhangs to allow ligation (Fig. 4-5). In such a way, the RFP cassette allows to make a selection of correctly assembled plasmids easier as transformed colonies which contained plasmids with added gRNAs were white and not red (Fig.4-4D).

Secondly, mRuby expression and PiggyBac integration was tested by transfecting NSCs with empty Ai1 plasmid plus PiggyBac transposase. Two weeks later, cells were analyzed using flow cytometry (Fig. 4-4E.). A proportion of cells were mRuby-positive, demonstrating that mRuby reporter and PiggyBac homology arms are functional.

Thirdly, functionality of the dCas9-VP160 part was also tested by comparing activation potential of empty Ai1 with plasmid containing only dCas9-VP160 (Fig. 4-4F). Activation was achievable using both plasmids to comparable levels. This empty Ai1 vector was used as the backbone for all multiplexing experiments that follows.

We tested how to efficiently clone multiple gRNAs into empty Ai1 plasmid by using the EMMA system. To achieve this, unique gRNA sequences were PCRed out

from corresponding gRNA-expression plasmids using PCR primers with unique overhangs. Those overhangs contained two restriction digestion sites (BsaI and BsmBI) on both sides (green and blue triangles in Fig. 4-5A respectively). This allowed us to firstly clone gRNAs into intermediate part vectors by BsaI-based GG reaction and then combine them all with empty Ai1 plasmid using second GG reaction based on BsmBI digestion enzyme. The resulting GG reaction products are transformed and correctly assembled plasmids are identified by color (Fig. 4-4B). This method allows adding multiple gRNAs into empty Ai1 vector which already contains dCas9-VP160.

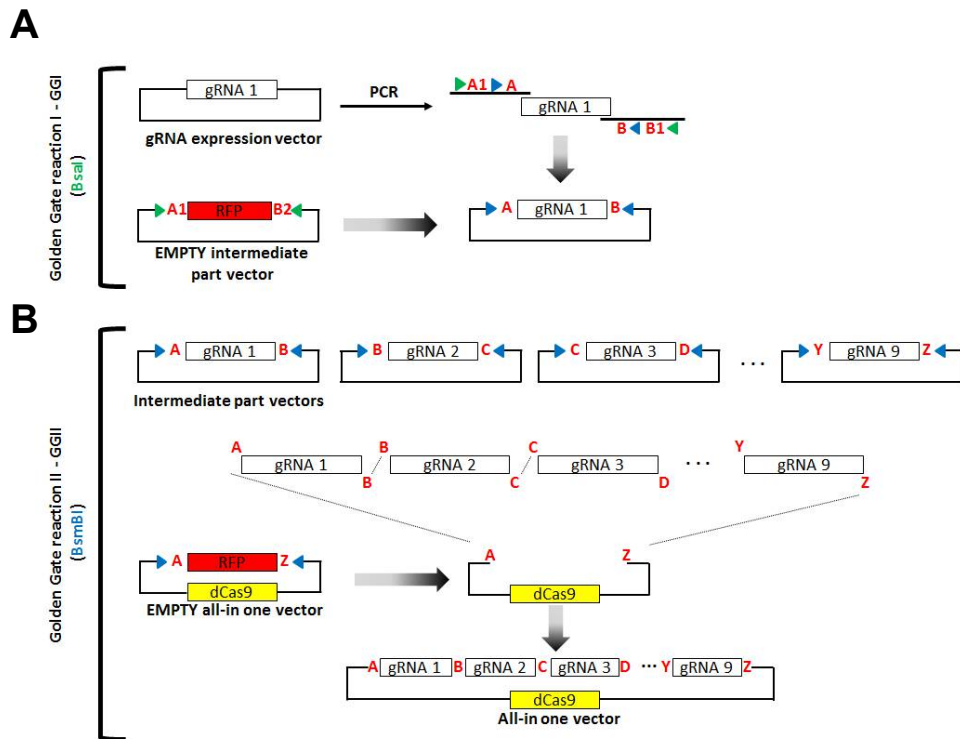


Figure 4-5 | Adaptation of EMMA system for assembly of plasmids containing multiple gRNAs and dCas9-VP160.

(A) Graphical representation of how gRNAs are cloned into intermediate part vectors using GG reaction. (B) Graphical representation of how multiple intermediate part entry vectors containing gRNAs are cloned into empty Ai1 vector using GG reaction.

One of the potential issues for constructing plasmids for multiplex gene regulation is repetitive sequences within such plasmids. Multiple identical U6 promoters could increase the chances of recombination either in *E.coli* during cloning or in mammalian cells after delivery, resulting in non-functional plasmids. To minimize such risks, we have decided to test a system which contains single U6 promoter and have tRNA-coding sequences separate multiple gRNAs (Fig. 4-6A). This strategy was inspired by the studies done in rice, however by the time this work was carried out there was no published methods for mammalian systems (Chun Wang et al. 2015; Dong et al. 2017).

Using method described in Figure 4-5, first plasmid for multiple gene regulation has been constructed and it was named tSON (Fig. 4-6A). This plasmid contains a single U6 promoter which drives expression of a transcript containing three gRNAs (targeting *Sox10*, *Olig2* and *Nkx6-2*) separated by tRNA sequences for Glu (Fig. 4-6B). After transcription, endogenous RNase P and RNase H cleaves this transcript near the ends of tRNA sequence, resulting in gRNA release from the transcript. We hypothesized that such gRNAs when can form complexes with dCas9-VP160 and activate gene transcription.

To test this hypothesis, we have transfected MEFs with either the tSON plasmid (“tSON”) or dCas9-VP160 and gRNA-expressing plasmids containing same gRNAs as ones’ present in tSON, namely g2, g17 and g30 (“Positive control”). mRNA fold change was measured using qPCR three days after transfection. It was clear that the gRNA-tRNA transcript was cleaved as intended as *Olig2* and *Nkx6-2* activation was achieved using tSON (Fig. 4-6C). *Sox10* activation was not obvious in either conditions. Interestingly, there was a small drop in activation potential (on average)

when using tSON plasmid compared to positive control. This most likely was due to the amount of each gRNA that was transcribed in different systems.

Although this strategy was evidently viable at small scale, we found when constructing larger plasmids, containing eight tRNA-gRNA, that it was difficult to build these vectors complicated. After numerous cloning attempts, no correctly assembled plasmids were found, assumingly because of difficulties to clone highly repetitive sequences. In the future, it might be possible to use different tRNAs sequences between gRNAs to reduce recombination risk but we have decided to construct plasmids which contained multiple U6 promoters instead.

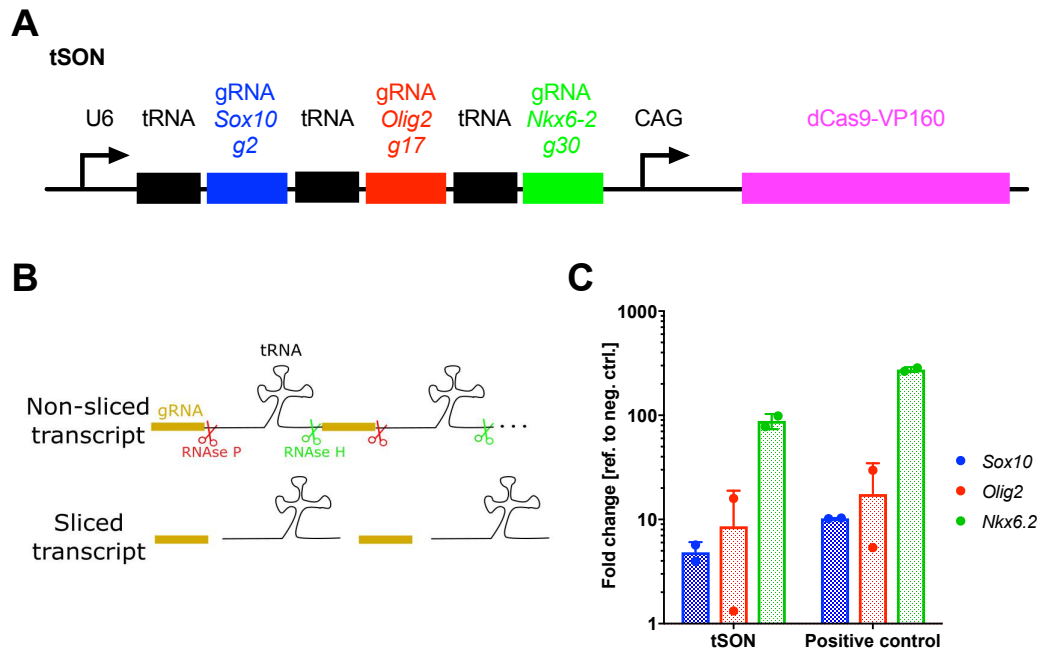


Figure 4-6 | U6-tRNA-gRNA-tRNA array system for multiplex gene regulation.

(A) Graphical representation of tSON plasmid. Three different gRNAs are added to the tSON, each targeting either *Sox10*, *Olig2* or *Nkx6-2*. These gRNAs are separated by sequences for Glu tRNA. (B) Graphical representation of how tRNA-gRNA system works. Single U6 promoter drives the expression of single transcript which contains tRNA-gRNA repeats. Such transcript is then spliced by RNase P and RNase H to release gRNA from tRNAs and become available for complex formation with dCas9-VP160. (C) qPCR result showing levels of *Sox10*, *Olig2* and *Nkx6-2* mRNA in MEFs three days after transfection with either tSON or dCas9-VP160 and g2, g17 and g30 gRNA-expressing plasmids (positive control), n=2.

To test gene regulation using plasmids with multiple U6 promoters, we have constructed three plasmids which contains increasing number of gRNAs targeting *Sox10*, *Olig2* and *Nkx6.2*. SON plasmid contained three gRNAs (one for each gene), 2SON contained five and 3SON contained eight gRNAs (Fig. 4-7A-C). Each gRNA is driven by a separate U6 promoter and proceeded by 30bp random sequence.

To test if these plasmids encoding sTFs are able to activate *Sox10*, *Olig2* and *Nkx6-2* we have transfected MEFs and measured mRNA fold change three days after. As predicted, all plasmids were capable of activating transcription of three different genes (Fig. 4-7D). Moreover increasing a number of gRNAs that target *Sox10* (SON to 3SON) increased the fold change that was observed. Interestingly, the fold change of *Nkx6-2* decreased with increasing size of the plasmid. One explanation might be that multiple proximal U6 promoters might partly interfere with each other resulting in decreased transcriptional activity in most distal U6 promoters. However, it is hard to determine the precise reason from this data as more than one variable varies (number of gRNAs, variations of gRNAs, size of the plasmid, number of U6 promoters). One thing we can conclude is that multiple genes can be activated in the same cell using single plasmid.

We re-transformed 2SON plasmid which contained five U6-gRNA units, picked six clones and sequenced them. All of the clones were correctly assembled, suggesting that recombination in bacteria is not an issue at this scale (Fig. 4-8). Furthermore, stably integrated plasmid containing eight U6-gRNA units remained functional throughout multiple cell divisions (Fig. 5-4 and Fig. 5-9), suggesting low risk of recombination in mammalian cells as well.

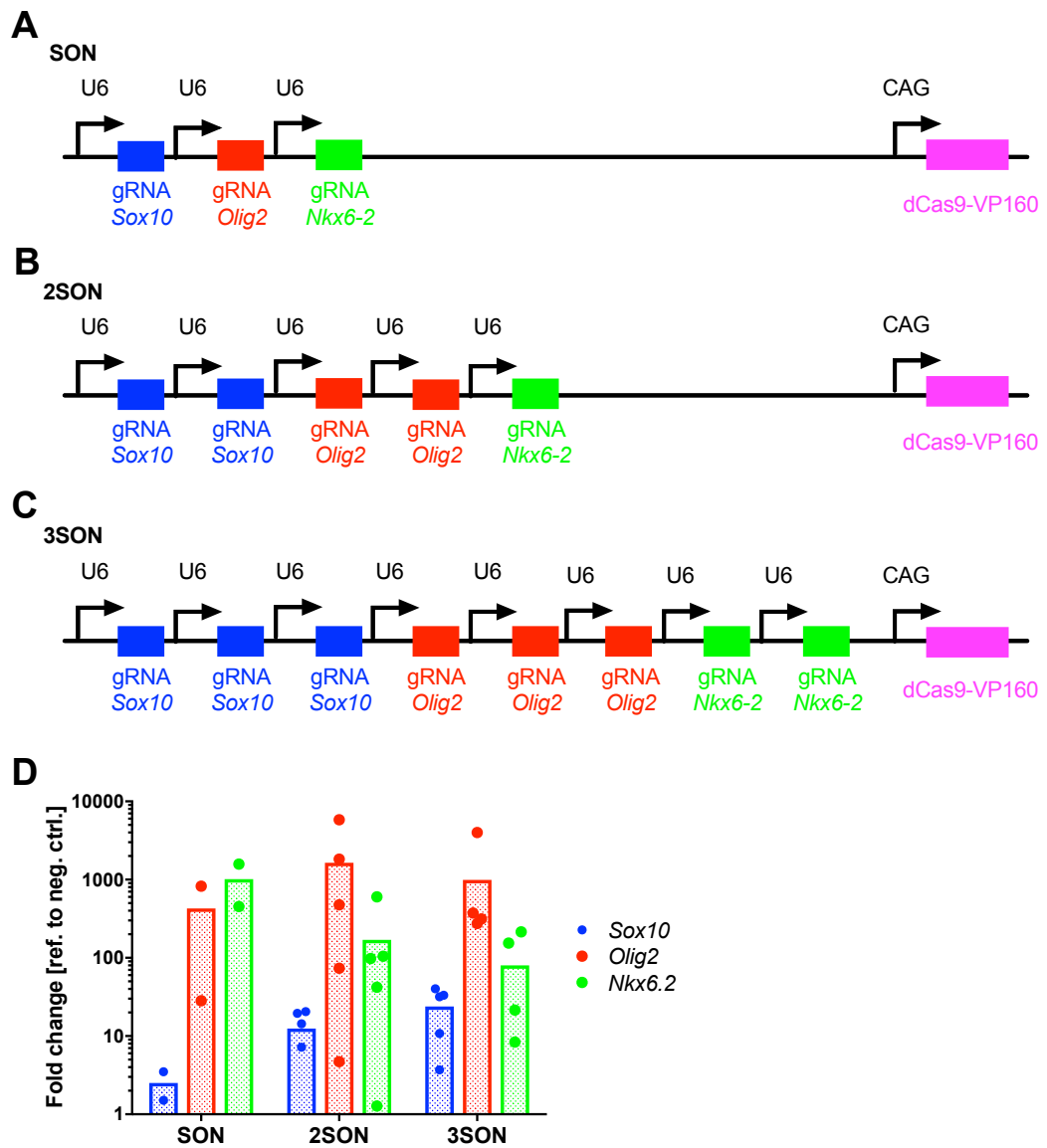


Figure 4-7 | U6-gRNA array system for multiplex gene regulation.

(A) Graphical representation of SON plasmid, containing three U6 promoters and three gRNAs targeting *Sox10*, *Olig2* and *Nkx6-2*. (B) Graphical representation of 2SON plasmid, containing five U6 promoters and five gRNAs targeting *Sox10*, *Olig2* and *Nkx6-2*. (C) Graphical representation of 3SON plasmid, containing eight U6 promoters and eight gRNAs targeting *Sox10*, *Olig2* and *Nkx6-2*. (D) qPCR result showing levels of *Sox10*, *Olig2* and *Nkx6-2* mRNA in MEFs three days after transfection with either SON, 2SON or 3SON, each dot represents biological replicate.

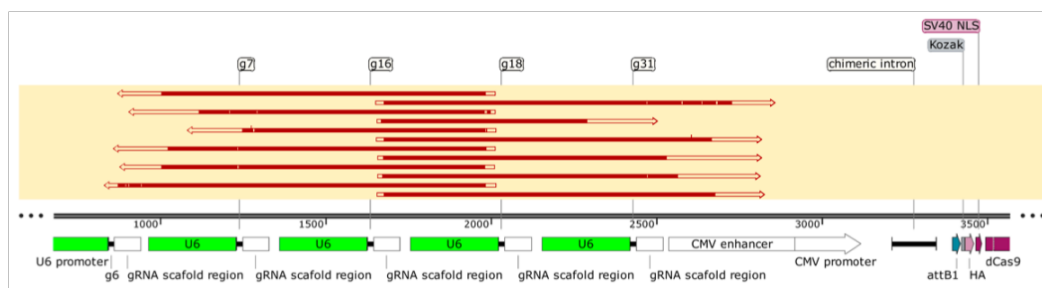


Figure 4-8 | Sequencing of 2SON clones after re-transformation.

Map and sequencing results of 2SON clones. 2SON has been re-transformed into competent *E. coli* bacteria and six clones were picked for sequencing to check for recombination events.

4.2 Construction and validation of system for simultaneous repression and activation

By constructing Ai1 plasmids and capability to activate multiple genes using them we have solved one of the multiplexing and delivery problems. However, the second problem still remains – how can we activate and repress different genes in the same cell at the same time? If cells were transfected with dCas9-VP160, dCas9-KRAB, some gRNAs targeting *Sox10* for activation and some gRNAs for *Sox9* repression there would be no activation or repression as different effectors would compete for the same promoters (Fig. 4-9A).

To overcome this problem, we have decided to adapt recently reported Cas9 mutation (Kleinstiver et al. 2015). This mutation, termed VRER, is four base pair substitution which results in four amino acid changes. As a result, such a Cas9 has altered PAM specificity. Instead of looking for usual NGG, VRER Cas9 mutant looks for NGCG PAM sequence. If such VRER mutant is to be used together with non-mutant Cas9, it is theoretically possible to target two genes exclusively by simply designing gRNAs which have different PAMs. If such mutation is to be made on dCas9, theoretically different gene transcription could be exclusively modulated at the same time (Fig. 4-9B).

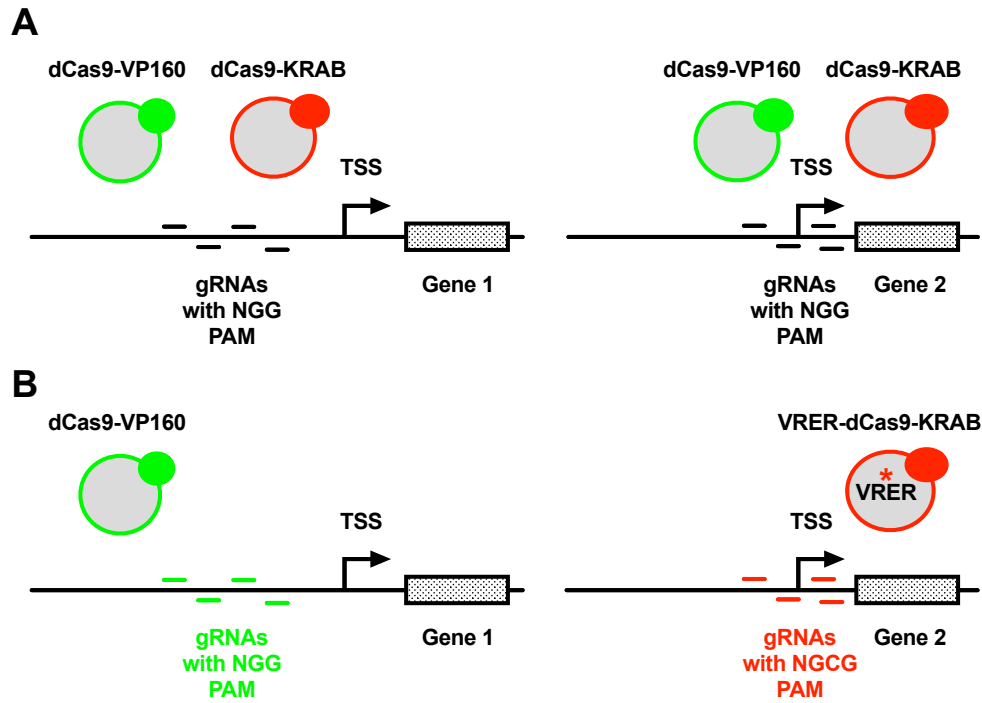


Figure 4-9 | System for simultaneous gene activation and repression.

(A) Graphical representation of current issue with simultaneous use of two opposing dCas9 proteins. As dCas9 proteins would recognize all gRNAs without specificity, dCas9-KRAB (red) and dCas9-VP160 (green) would bind promoters of both, genes which are intended for activation (left side) and for repression (right side). (B) Graphical representation of the system where dCas9 proteins recognize different gRNAs and therefore effectors can be targeted to different promoters in exclusive manner. This is achieved by mutating dCas9-KRAB (red, VRER) and adjusting its PAM specificity to NGCG, while dCas9-VP160 will continue to target NGG PAM.

4.2.1 Construction and validation of VRER dCas9 mutants for altered PAM specificity

To insert VRER mutation in dCas9 protein, we have used site-specific mutagenesis (Fig. 4-10A). This method is based on introducing a small DNA change into a plasmid by using PCR primers which have mutated sequence (Fig. 4-10B). The PCR product is then phosphorylated and ligated. Importantly, all this is done in the same reaction tube which means original non-mutated plasmid which was used as PCR template is still present. To remove original (non-mutated) plasmid, sample is treated with DpnI restriction enzyme which cleaves only methylated sites. In such way, because PCR products are not methylated while plasmids after transformation are, only original plasmid is digested into pieces. This is followed by transformation and sequencing of the plasmids to confirm new mutation.

After four rounds of site-specific mutagenesis, all four mutations were introduced into dCas9 plasmid (Fig. 4-10C).

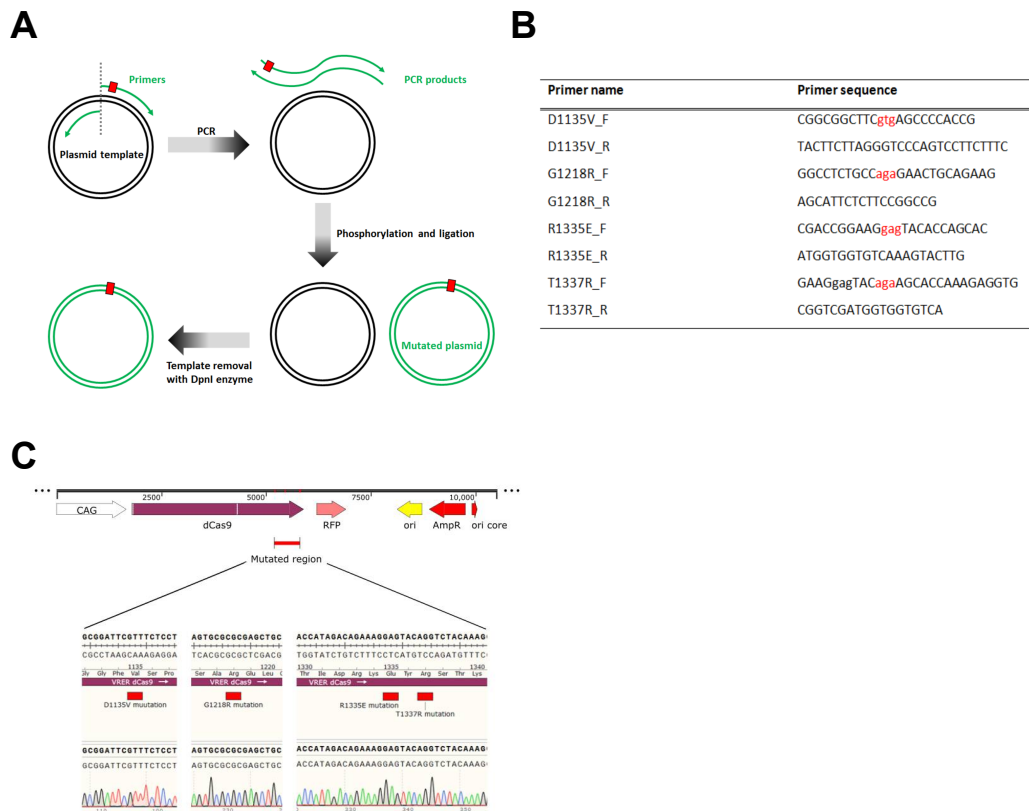


Figure 4-10 | Making VRER dCas9 mutants.

(A) Graphical representation of the site-specific mutagenesis, which was used to create four amino acid substitutions in dCas9 protein. (B) Table outlining primers and their sequences which have been used to introduce mutations in dCas9 via site-specific mutagenesis. (C) Map and sequencing results of VRER dCas9 mutant plasmid demonstrating that four amino acid substitutions (VRER) have been successfully introduced into dCas9 plasmid.

To check if VRER mutations would make dCas9 specific to gRNAs with NGCG PAM, we needed two sets of gRNAs targeting same promoter – one with NGG and one with NGCG PAMs. We already had gRNAs that can activate *Sox10* in NSC and these gRNAs have NGG PAM (green boxes in Fig. 4-11A). Additionally, we have designed another set of gRNAs around the same location, but these gRNAs have NGCG PAM (red boxes in Fig. 4-11A). PG1.1 were transfected with either dCas9-VP160 or mutant VRER-dCas9-VP160 and either gRNAs with NGG PAM (g1-g10) or gRNAs with NGCG PAM (g153-g157). Three days after transfection, mRNA levels were measured (Fig. 4-11B). From the data, it is clear that NGCG gRNAs do not work with dCas9-VP160 while NGG gRNAs do. However, neither gRNAs worked with VRER-dCas9-VP160. This might be because NGCG gRNAs were not functional as even slight location change can have a dramatic effect on gRNA potency.

As our second gene for VRER validation, we have chosen *Nkx6-2*. Just as for *Sox10*, we have designed new gRNAs with NGCG PAM for *Nkx6-2* (Fig. 4-11C). MEFs were transfected with either NGG or NGCG gRNAs and either dCas9-VP160 or VRER-dCas9-VP160. Three days after transfection cells were harvested and *Nkx6-2* mRNA fold change was measured. As predicted, VRER-dCas9-VP160 was able to activate *Nkx6-2* transcription only if NGCG gRNAs, but not NGG gRNAs, were used (Fig. 4-11D). To large extent, opposite was true for dCas9-VP160. However, a small level of activation was seen when using dCas9-VP160 with NGCG gRNAs (17-to-30 fold change). Although this is significantly lower compared to NGG gRNAs (249-to-4755 fold change), but it is still evident. Such findings suggest that although PAM specificity is strongly favored but some overlap might still exist between gRNAs with NGCG and NGG PAM.

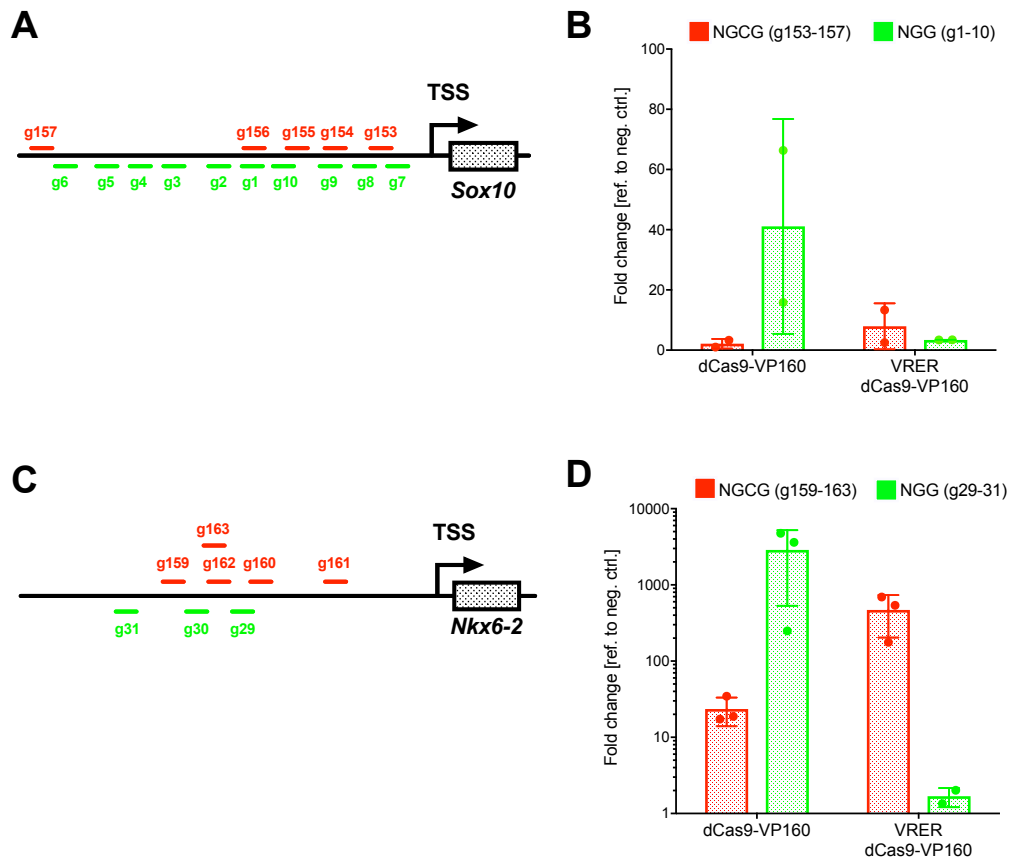


Figure 4-11 | Validating VRER exclusivity using dCas9-VP160.

(A) gRNA positioning along *Sox10* proximal promoter and transcription start site (TSS) for transcription activation. g153-g157 (red) have PAM NGCG which should be specific to VRER-dCas9, g1-g10 (green) have PAM NGG which should be specific to dCas9. (B) qPCR results showing *Sox10* mRNA levels in PG1.1 NSC three days after transfection with either dCas9-VP160 or VRER-dCas9-VP16 and either gRNAs g1-g10 or g153-g157, n=2. (C) gRNA positioning along *Nkx6-2* proximal promoter and transcription start site (TSS) for transcription activation. g159-g163 (red) have PAM NGCG which should be specific to VRER-dCas9, g29-g31 (green) have PAM NGG which should be specific to dCas9. (D) qPCR results showing *Nkx6-2* mRNA levels in MEFs three days after transfection with either dCas9-VP160 or VRER-dCas9-VP16 and either gRNAs g29-g31 or g159-g163, n=3.

4.2.2 Mouse gene repression in NSC using VRER-dCas9-KRAB

After validating the VRER system, we went further to create VRER-dCas9-KRAB mutants and investigate if such dCas9 can repress genes. Firstly, we focused on *Sox9* repression as we have already achieved it using non-mutant dCas9-KRAB (Fig. 3-13).

For this purpose, we have designed gRNAs for proximal *Sox9* promoter which had NGCG PAM (Fig. 4-12A). To test if these gRNAs could repress *Sox9*, SOX9-V5 NSCs were transfected with VRER-dCas9-KRAB and a pool of four gRNAs (g94-g97). Four days after transfection, cells were analyzed by immunocytochemistry and qPCR.

As observed previously, we could not see any significant change at mRNA level (Fig. 4-12E). However, more SOX9-low/negative cells were found after immunostaining (Fig. 4-12C and D). Approximately 25-8% more cells were SOX9-low based on three independent biological replicates.

To gain more confidence in the VRER-dCas9-KRAB ability to repress mammalian genes, we have conducted experiments to repress *Olig2* in PG1.1 NSC line. Same methodology was used as for *Sox9* repression and similar results were found. Although repression was not evident at mRNA level (Fig. 4-13F), more OLIG2-low/negative cells were observed after transfection with VRER-dCas9-KRAB and a pool of gRNAs (Fig. 13E).

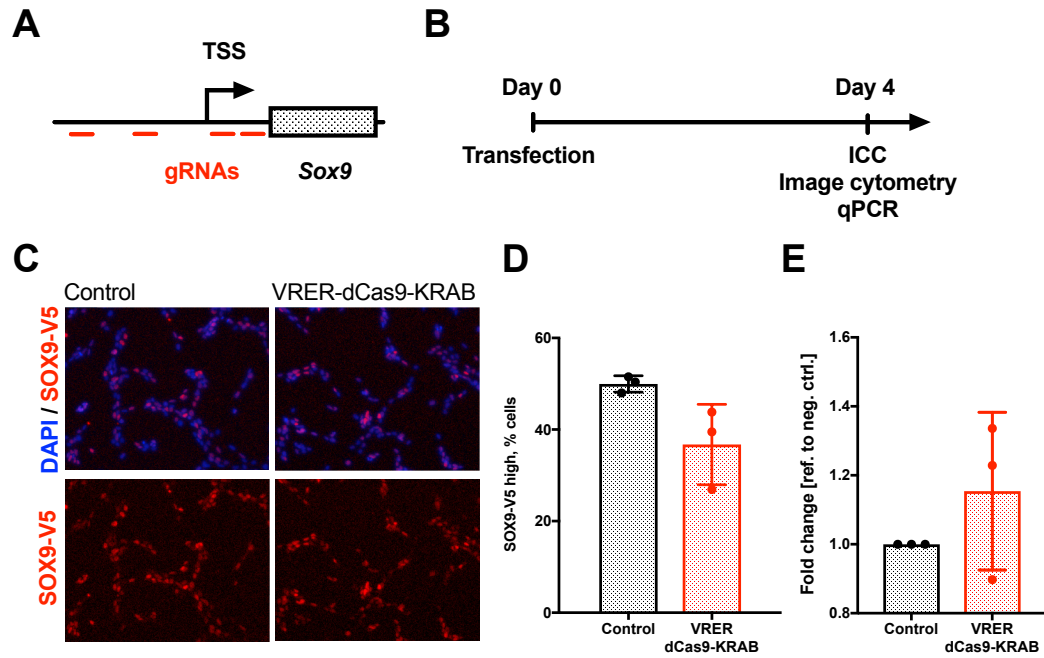


Figure 4-12 | *Sox9* repression using VRER-dCas9-KRAB in SOX9-V5 NSC line.

(A) gRNA positioning along *Sox9* transcription start site (TSS) for transcription repression. (B) Graphical representation of experimental setup. (C) Immunocytochemistry for V5 in SOX9-V5 cells after transfection with VRER-dCas9-KRAB and 4 gRNAs (g94-g97) targeting *Sox9* transcription start site (“VRER-dCas9-KRAB”) or non-transfected cells (“Control”). (D) Quantification of image cytometry data showing percentage of SOX9-V5 high cells after transfection with VRER-dCas9-KRAB and 4 gRNAs targeting *Sox9* (“VRER-dCas9-KRAB”) or in non-transfected cell (“Control”), n=3. (E) qPCR data of *Sox9* transcription levels four days after cells were transfected with VRER-dCas9-KRAB and 4 gRNAs targeting *Sox9* (“VRER-dCas9-KRAB”) or non-transfected cells (“Control”), n = 1.

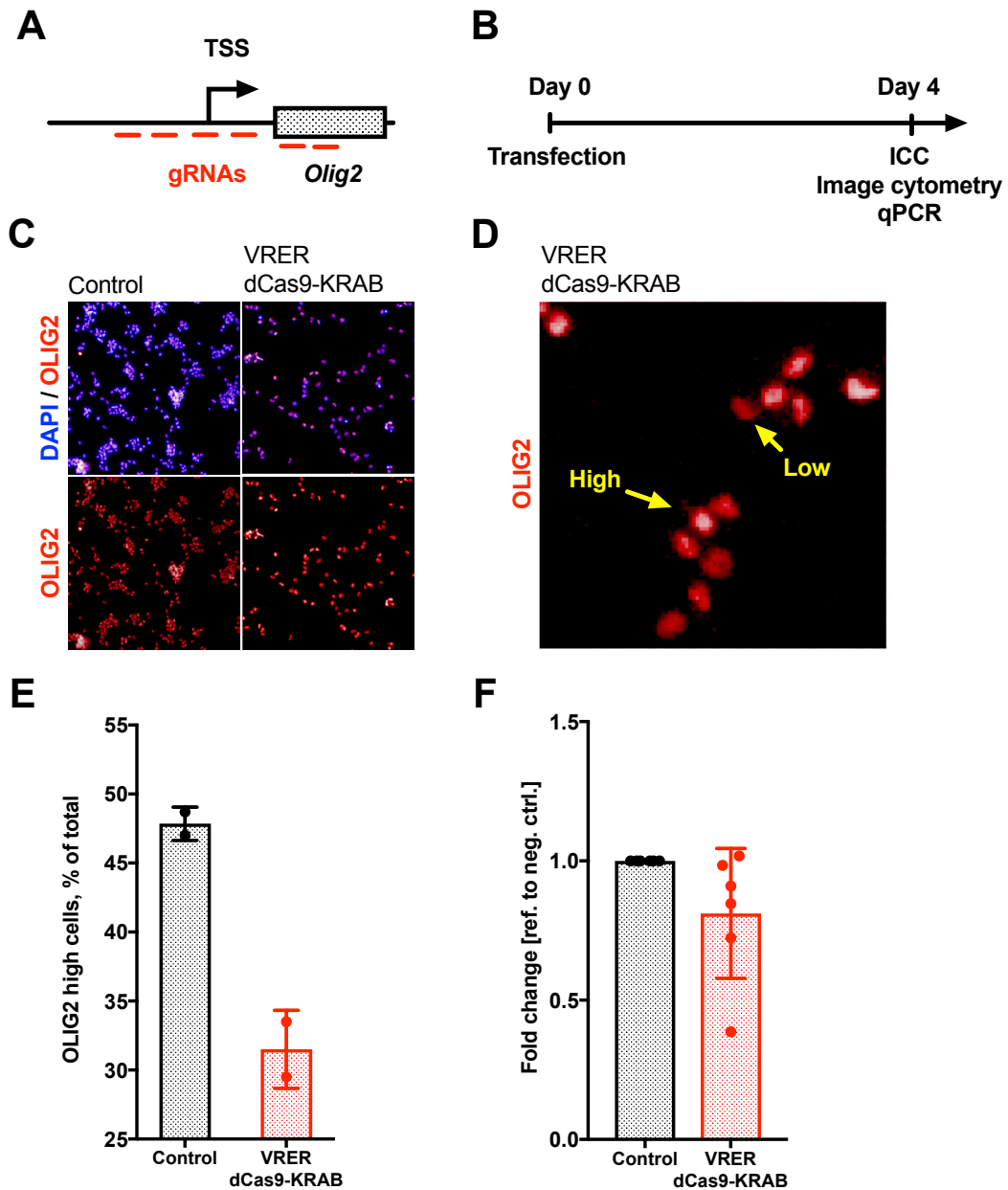


Figure 4-13 | *Olig2* repression using VRER-dCas9-KRAB in PG1.1 cell line.

(A) gRNA positioning along *Olig2* transcription start site (TSS) for transcription repression. (B) Graphical representation of experimental setup. (C) Immunocytochemistry for OLIG2 in PG1.1 cells after transfection with VRER-dCas9-KRAB and 6 gRNAs (g175-g180) targeting *Olig2* transcription start site ("VRER-dCas9-KRAB") or non-transfected cells ("Control"). (D) Image of stained OLIG2 in PG1.1 cells to demonstrate gating strategy based on OLIG2 intensity. Examples of low and high OLIG2 cells are shown with yellow arrows. (E) Quantification of image cytometry data showing percentage of OLIG2 high cells after transfection with VRER-dCas9-KRAB and 6 gRNAs targeting *Olig2* ("VRER-dCas9-KRAB") or in non-transfected cell ("Control"), n=2. (F) qPCR data of *Olig2* transcription levels four days after cells were transfected with VRER-dCas9-KRAB and 6 gRNAs targeting *Olig2* ("VRER-dCas9-KRAB") or non-transfected cells ("Control"), n = 5.

4.3 Construction of a *Sox10*-V5 tagged NSC line for multiplexing validation

4.3.1 Evaluating SOX10 antibody specificity in our cell lines

To test the multiplex capability of activating and repressing genes at the same time (in the same cell), we were required to find antibody for SOX10 that works in our system. We have tested a number of different antibodies which have been reported to work in other cell lines (Fig. 4-14A). However, after testing these antibodies in differentiated PG1.1 NSC, we could not see any staining that was not a background (Fig. 4-14B-C). Importantly, we knew that differentiating PG1.1 NSC line had increased *Sox10* mRNA levels. Therefore, we have concluded that none of the antibodies are functional. To overcome this hurdle, we decided to epitope tag *Sox10* gene with V5 tag knock-in using CRISPR.

4.3.2 Tagging *Sox10* with C-terminus V5 in PG1.1 and BL6 NSC lines

None of the antibodies that we tested worked on PG1.1 cell line. Therefore, we decided to tag *Sox10* with V5 to allow SOX10 protein visualization. To achieve this, we have used Cas9 protein, gRNAs targeting *Sox10* C-terminus and single-stranded donor DNA (ssODN) which contained sequence V5 tag. Tagging works based on the following principle. After transfecting cells with all three components, Cas9 and gRNAs cause double-strand break (DSB) which is then repaired with a help of ssODN. As a result, the gene of interest is tagged with V5 peptide at the end of C-terminus to create a fusion protein product.

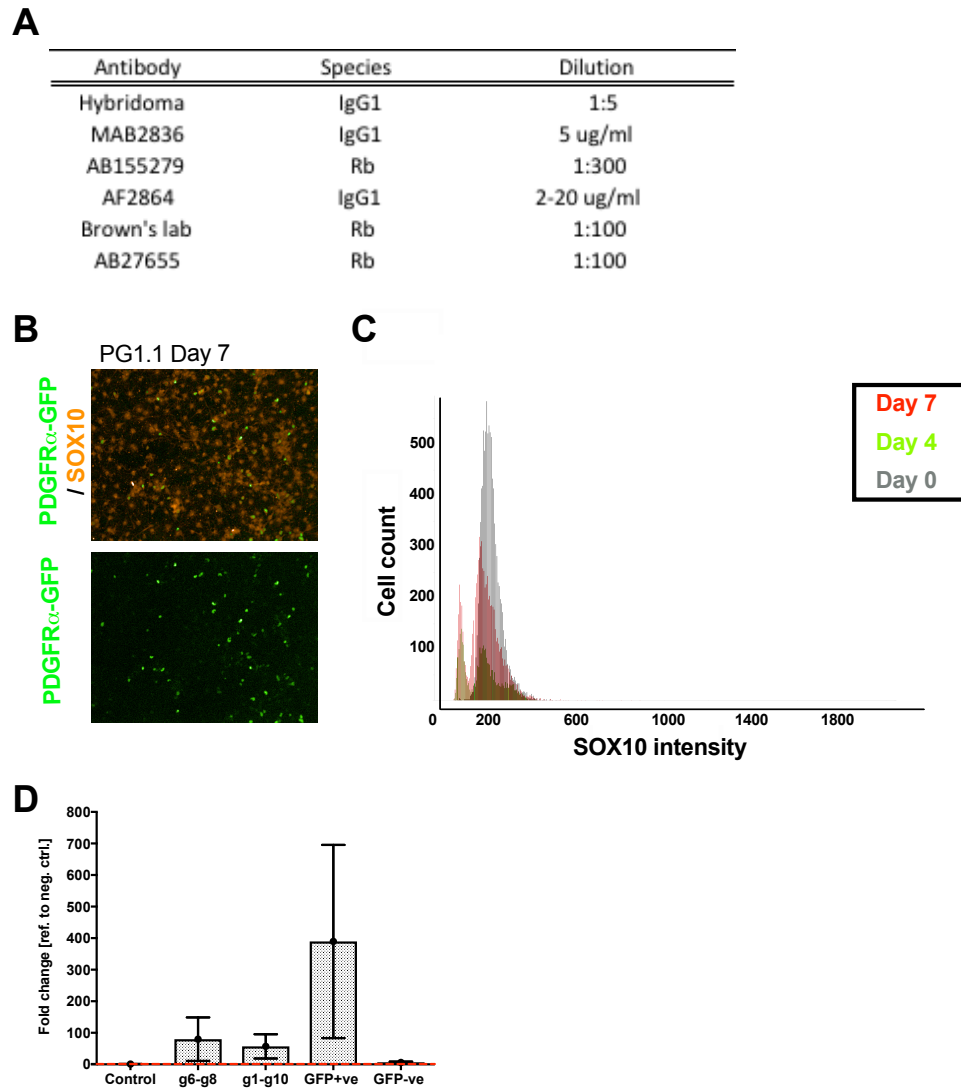


Figure 4-14 | Evaluating SOX10 antibodies.

(A) Table summarizing tested anti-SOX10 antibodies. (B) Immunocytochemistry of SOX10 and PDGFR α -GFP seven days after differentiation. AB155279 antibody was used in this example. (C) Image cytometry analysis of SOX10 immunocytochemistry in self-renewal conditions (grey), four days after differentiation (green) or seven days after differentiation (red). AB155279 antibody was used in this example. (D) qPCR results showing mRNA levels of *Sox10* either after activation with dCas9-VP160 and specified gRNAs (g6-g8 and g1-g10) or of differentiated PG1.1 GFP-positive and GFP-negative NSCs (after sorting), n=1.

To validate if gene targeting was successful, genotyping was performed using the strategy outlined in Figure 4-15B. After genotyping transfected PG1.1 cells, a band was observed at expected 600bp size while parental line (non-transfected PG1.1 NSC) did not produce band of such size (Fig. 4-15C). To validate SOX10 tagging using immunocytochemistry, we differentiated tagged PG1.1 cells to OPCs and stained for V5 and PDGFR α -GFP (control for differentiation). Although cells differentiated and became PDGFR α -GFP positive, no V5-positive cells could be identified (Fig. 4-15D). We reasoned that derivation of clonal lines might be necessary as if only low percentage of cells were tagged and only 20% of those differentiated, it might be hard to find positive cells without deriving clonal lines.

We have seeded 1×10^5 PG1.1 SOX10-V5 cells into 10cm² dish and cultured them for 7-12 days. Colonies which grew were picked and plated into 96-well plate for further expansion. After we had enough cells, genomic DNA was extracted from eight different clonal lines and genotyping was performed. Two out of eight clonal lines were positive (Fig. 4-16A), which gives an estimated tagging efficiency of 25%. These two clonal lines (C2 and C7) showed normal NSC morphology (Fig. 4-16B). However, no V5-positive cells were found after differentiation of either of the clonal lines (Fig. 4-16C). To continue our work, we decided to test an independent NSC line called Black 6 (BL6), which can more efficiently differentiate into mature *Sox10* expressing cells (Fig. 5-8).

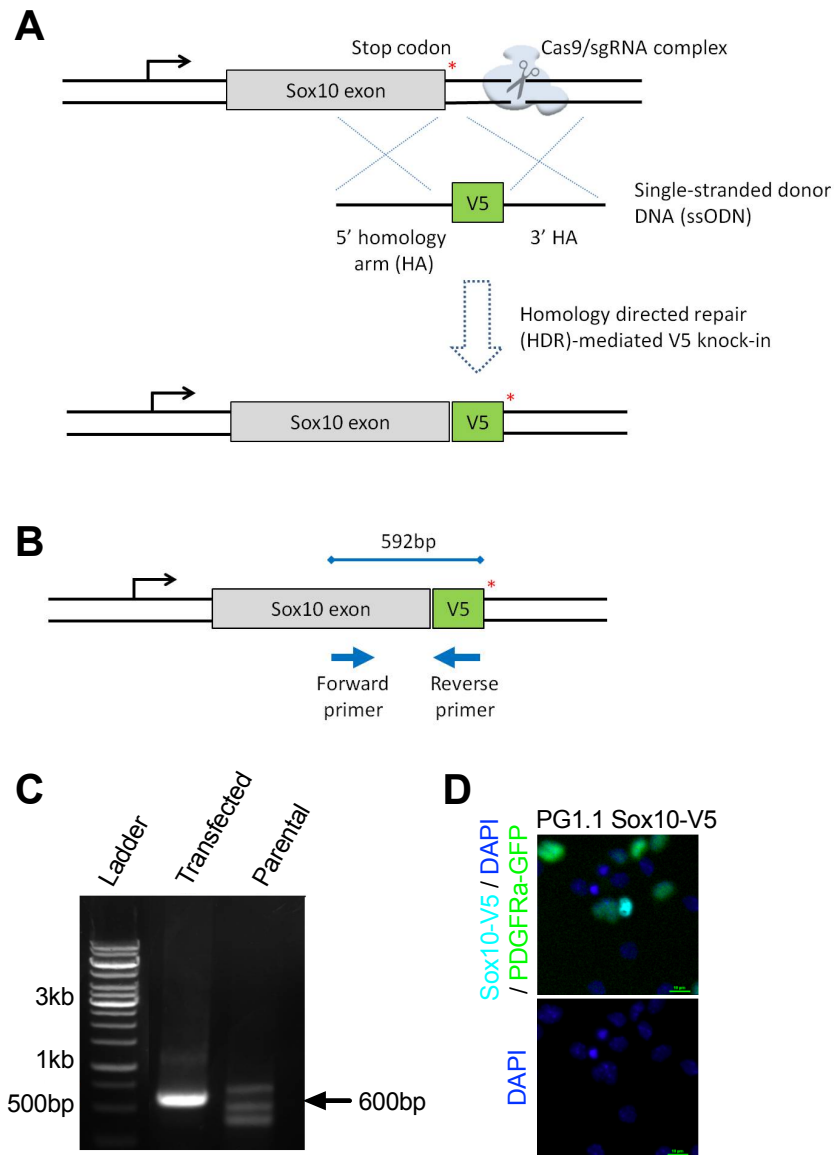


Figure 4-15 | Strategy for V5 tagging of C-terminus of *Sox10* using CRISPR-Cas9 protein and in vitro transcribed gRNAs in PG1.1 NSC line.

(A) Graphical representation of the process of V5 tagging, which involves cutting near *Sox10* stop codon to initiate homology directed repair (HDR)-mediated insertion of single strand donor containing V5 tag sequence. (B) Genotyping strategy. One primer is placed inside *Sox10* exon and second one is placed on V5. If tagging was successful, genotyping using these primers and DNA from treated cells as template will generate 592bp band on agarose gel. (C) Genotyping results of PG1.1 NSC which were tagged with V5 at the C-terminus of the *Sox10*. Parental (non-treated) cells were used as negative control. (D) Immunocytochemistry for V5 and PDGFR α -GFP of tagged and differentiated PG1.1 NSC.

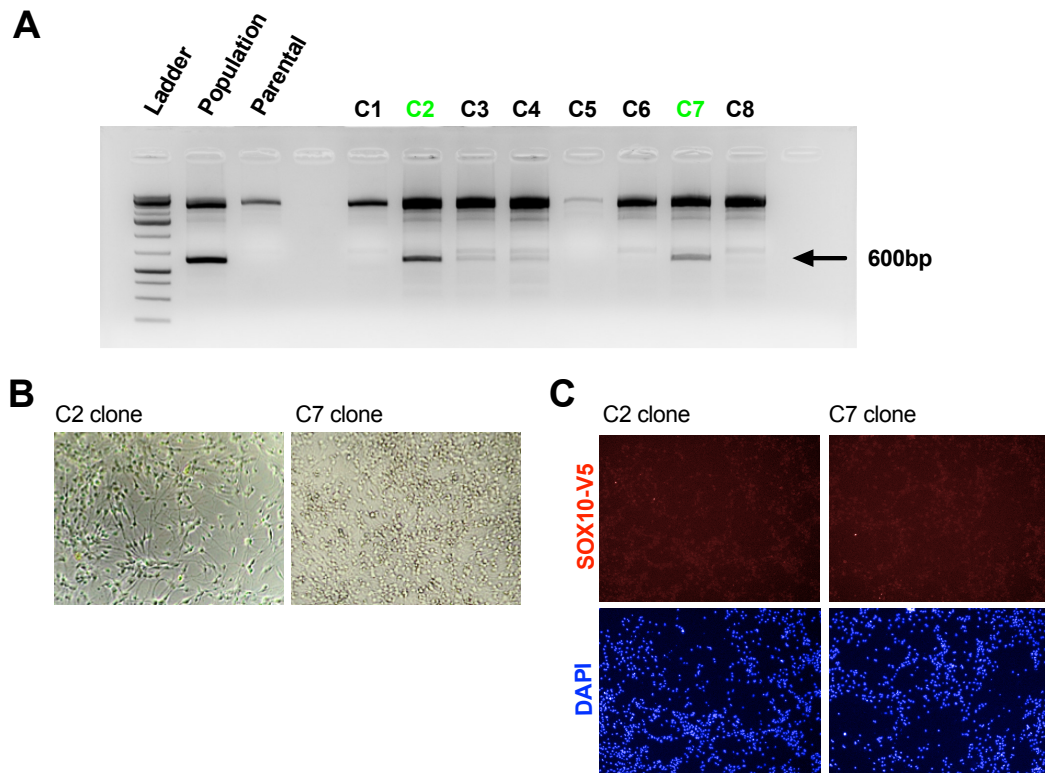


Figure 4-16 | Derivation of PG1.1 SOX10-V5 clonal cell lines.

(A) Genotyping of clonal PG1.1 SOX10-v5 cell lines (C1 to C8). “Population” is population of PG1.1 SOX10-V5 cells (positive control) and “Parental” is parental PG1.1 NSC line (negative control). (B) Phase contrast picture demonstrating NSC morphology of clone 2 (C2) and clone 7 (C7). (C) Immunocytochemistry of PG1.1 SOX10-V5 C2 and C7 clones after differentiation.

Using the same experimental strategy described above, we tagged *Sox10* in BL6 NSCs with a V5 tag. We have used two different gRNA sets (#1 and #2) and both gave cell populations which were genotyped positively for the V5 tag (Fig. 4-17A). After forcing these two populations to differentiate into OPCs, we could only find a couple V5-positive cells in population #1 and no V5-positive cells in population #2 (Fig. 4-17C). After picking and genotyping clones, we found two clonal lines (C6 from population #1 and C4 from population #2) to be successfully tagged (Fig. 4-17B). Importantly, after inducing these clones to differentiate, a proportion of C6 cells were observed to be V5-positive, while clone C4 did not have any V5 positive cells. Such result could potentially be explained by difference in capability to differentiate between clonal lines or by some genomic abnormalities that were caused at *Sox10* locus during tagging process when using #2 gRNA set.

The BL6 SOX10-V5 C6 clonal line will be used in the future studies for multiplexing experiments to test if multiple genes can be activated and repressed in the same cell using systems that we have built; but time constraints meant we did not pursue it further in this thesis.

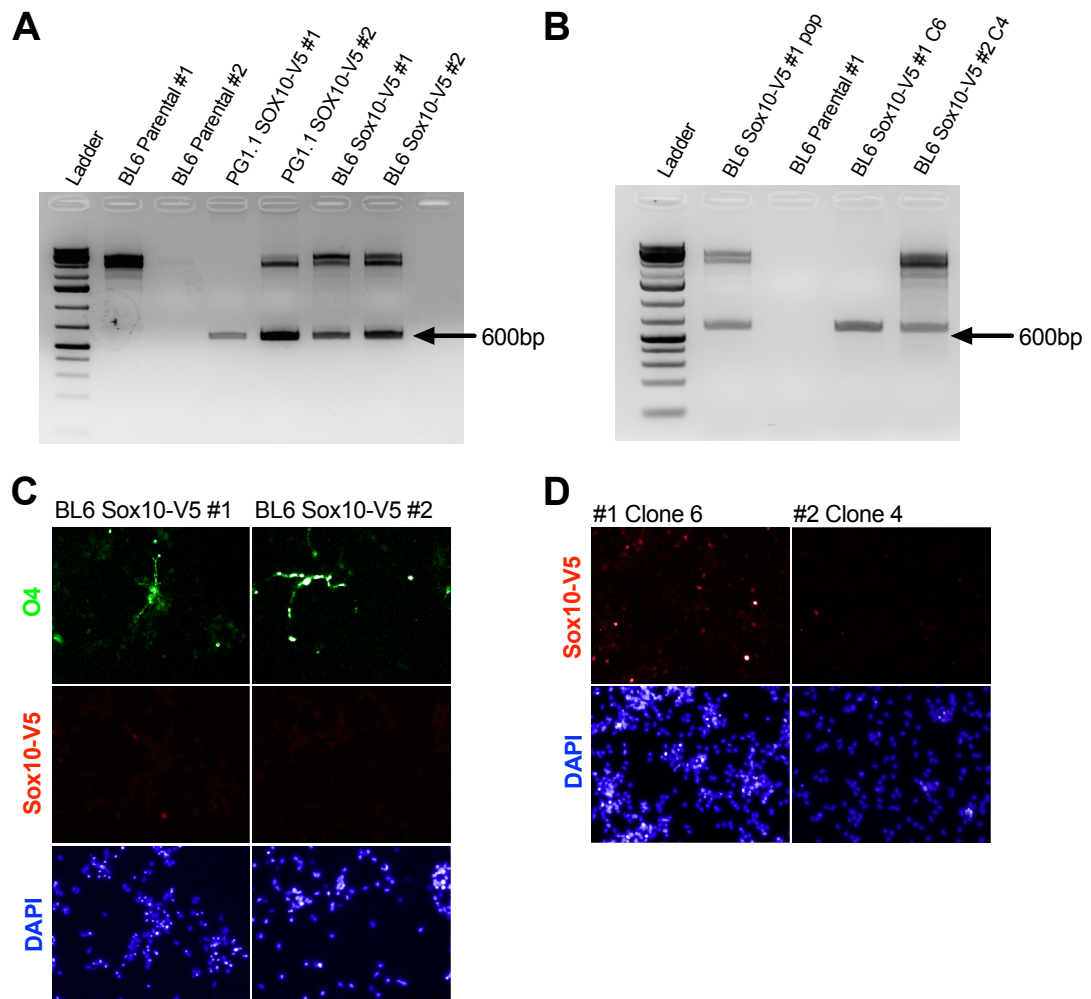


Figure 4-17 | Derivation of BL6 SOX10-V5 clonal cell lines.

(A) Genotyping of tagged BL6 cells. Pop#1 and Pop#2 are two populations of BL6 SOX10-V5 cells (tagged using two different gRNAs), “Parental” is parental BL6 NSC line (negative control) and “PG1.1 SOX10-V5” are tagged populations of PG1.1 SOX10-V5 cells (positive control). (B) Genotyping of clonal BL6 SOX10-V5 cells. (C) Immunocytochemistry of differentiated BL6 SOX10-V5 tagged NSC populations (OL stage). (D) Immunocytochemistry of differentiated BL6 SOX10-V5 tagged clonal lines (OPC stage).

4.4 Concluding remarks

Interesting findings and lessons from the work in Chapter 4 will be highlighted in the following paragraphs. Firstly, we have successfully built an easy-to-use and rapid system for adding multiple gRNAs into the plasmid with dCas9-V160. Up to eight U6-gRNA units (highly repetitive sequence) have been added in an array without any evidence of recombination in *E.coli* or mammalian cells. By the time this work was performed, only six gRNAs have been reported to be successfully multiplexed using single plasmid. Importantly, using this all-in-one plasmid, we have successfully activated three different genes (*Sox10*, *Olig2* and *Nkx6-2*) from a single vector in MEFs. This was significant improvement over present state-of-art methods, as it was more modular (easier manufacturing with possibility to expand a number of gRNAs used) and included dCas9 in the same vector.

Secondly, we have successfully built the capability for directing dCas9 proteins to different gRNAs (and therefore promoters) in almost exclusive manner. This was achieved by introducing four mutations into dCas9 protein. Such mutant dCas9, called VRER dCas9, has been reported to have a new PAM specificity (NGCG instead of NGG). We have validated strong preference of VRER-dCas9 to NGCG PAM, however non-mutant dCas9 also recognizes NGCG PAM to some extent. Importantly, using VRER-dCas9-KRAB we have successfully repressed *SOX9* and *OLIG2* in NSC lines.

Thirdly, we have built Sox10-V5 NSC cell lines for easier SOX10 detection. This was necessary because none of the commercial antibodies against SOX10 were functional in our hands. As a result, we have generated PG1.1 and BL6 NSC lines

which had SOX10 tagged with V5 based on the genotyping results. We have identified BL6 NSC clonal cell line which were successfully tagged, as confirmed by genotyping and V5 staining upon induction of differentiation. Interestingly, we have also observed multiple times that it is misleading to rely solely on genotyping results, as a number of clones had correct bands in genotyping but were demonstrated to be negative by immunostaining.

5 Programing and reprograming cell identity using sTFs

5.1 Directing NS cell differentiation by overexpression or sTF-driven activation of key OPC master regulators

The preceding chapters have demonstrated biochemical activity and transcriptional activation of target genes can be achieved in NS cells and MEFs for some of the key OPC ‘master regulators’, such as *Sox10*, *Olig2* and *Nkx6-2*. In this chapter, we explore whether activation of these genes can be sufficient to direct cell fate. We explore two different contexts. First, the ‘programing’ of differentiation of NS cells to see if OPCs/OL can be made at increased efficiencies. i.e. can we control the differentiation pathways. Second, can we directly reprogram MEFs into OPCs?

5.1.1 Construction of PG1.1 cell line with inducible SOX10 and OLIG2 overexpression (PG1.1-SO)

To test if transcription activation and overexpression can force NSCs to differentiate into OPCs and OLs, we have made use of the OPC reporter NSC cell line PG1.1 (described in Chapter 3). Following transfer to differentiation conditions (EGF withdrawal; plus FGF, PDGF-AA and Forskolin), ~20% of cells become GFP positive. These GFP positive cell are early OPCs, as the H2B-GFP cassette is driven by the endogenous PDGFR α promoter. Neurons, astrocytes and neural stem cells do not express PDGFR α . Unless otherwise stated, this cell line was used in the following experiments. SOX10 and OLIG2 are two of the critical reprograming factors that have been described in past work using cDNA. As a positive control for our sTF experiments we created a new NS cell line that enabled inducible overexpression of the cDNA encoding these two factors.

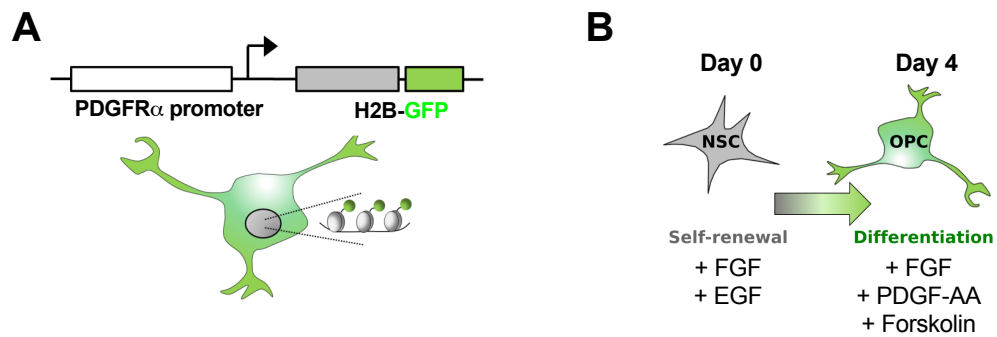


Figure 5-1 | PDGFR α -GFP (PG1.1) reporter cell line.

(A) Graphical representation of PG1.1 reporter mouse NS cell line. The PG1.1 cell line was constructed by Dr. Bartłomiej Baranowski (a previous PhD student in the Pollard lab). (B) Under self-renewal conditions (presence of FGF and EGF) PG1.1 cells lack any GFP fluorescence; however, upon differentiation (presence of FGF, PDGF-AA and Forskolin) cells with upregulated PDGFR α becomes GFP positive. This provides a convenient reporter to assess oligodendrocyte lineage specification.

This was achieved by constructing plasmid M1, which drives expression of *SOX10* and *OLIG2* after addition of DOX (using the TETON 3G system). This plasmid also harbours PiggyBac recombination sites (inverted repeats) and a blasticidin resistance cassette (Fig. 5-2 A-B). To generate a stable cell line, we transfected PG1.1 NSCs with the M1, PB_CAG_Tet3G plasmid together with the CAG_Pbase plasmid (encoding the transposase). Cells were cultured for two weeks in the presence of 1 μ g/ml blasticidin to select for positive cells (Fig. 5-2C). The resulting pool of transfected cells has been named PG1.1-SO.

To validate that PG1.1-SO cells are co-expressing *SOX10* and *OLIG2* after addition of DOX, we have conducted DOX dose response experiments over a time-course and performed qPCR for the transgenes (Fig. 5-2D). As it is evident from the data, increasing concentrations of DOX indeed triggers increased levels of *SOX10* and *OLIG2* expression from these cells. At Day 1 expression levels are low (100-fold over no DOX control), as transcriptional and translational processes did not have enough time to reach maximum rate. By Day 2 we see higher levels of expression (200 to 500-fold over no DOX control), while expression at Day 3 is highest (1500 to 7500-fold over no DOX control). However, the PG1.1-SO cell line was leaky, as the transgene was expressed even without addition of DOX (data not shown). To minimize leakage and therefore potential side-effects of the leakage, we needed to establish clonal cell lines from PG1.1-SO.

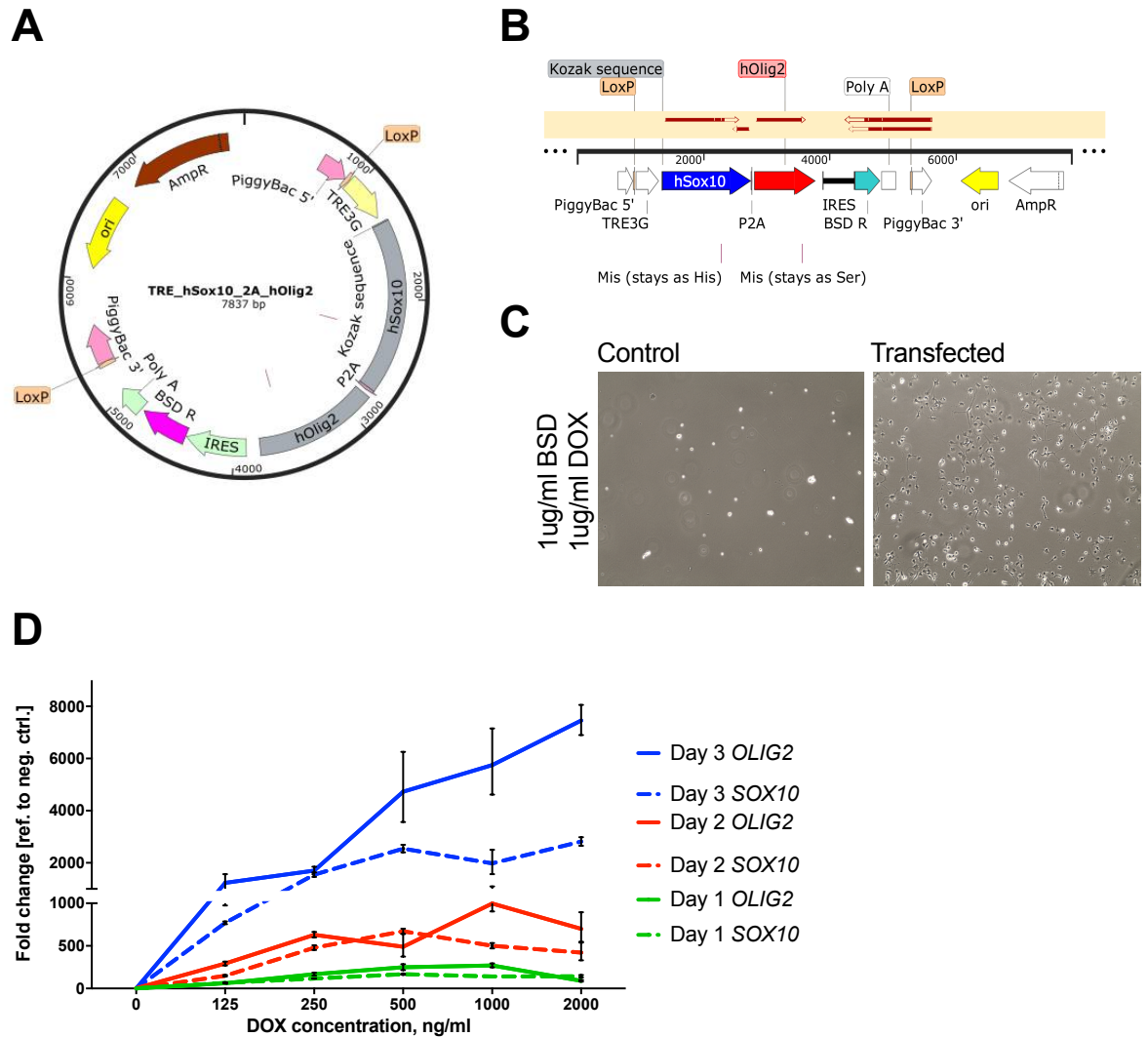


Figure 5-2 | Construction of PG1.1 cell line with DOX inducible *SOX10* and *OLIG2* (PG1.1-SO).

(A) Map of constructed M1 plasmid which was used to generate the DOX-inducible *SOX10* and *OLIG2* cell line (PG1.1-SO). (B) Sanger sequencing confirmed correct assembly of M1 plasmid. (C) Phase contrast images demonstrating survival of PG1.1-SO cells in presence of 1 µg/ml DOX after selection with 1 µg/ml blasticidin, while control non-transfected cells were killed. Prior to selection, PG1.1 cells were transfected with Tre_*SOX10*_p2A_*OLIG2* (M1), CAG_pBase and Pb_CAG_Tet3G plasmids and cultured for two weeks. (D) qPCR data showing mRNA levels of *SOX10* and *OLIG2* after PG1.1-SO cells were exposed to different DOX concentration. Results were normalized to 0 DOX control. RNA was collected at Day 1, Day 2 and Day 3, n=1.

To develop cell lines with minimal leaky expression we generated four clonal PG1.1-SO cell lines and tested *SOX10* mRNA by qPCR without adding DOX (Fig. 5-3A). Some of the clonal PG1.1-SO lines had high levels of *SOX10* expression even without DOX (>504-fold in c1 and >5572-fold in c3 over parental cell line) while in other cell lines there was minimal leakage (>183-fold in c2 and >71-fold in c4 over parental cell line). We decided to use PG1.1-SO c4 clonal line for further experiments as this had lowest leakage levels.

It is well known that differentiation can be triggered at high cell densities – even in the presence of EGF. To ensure that the effect on differentiation that we will be investigating is from overexpression and not due to secondary effects of enhanced proliferation/density, we have compared PG1.1-SO c4 growth rates to that of parental PG1.1 cell line with and without DOX (Fig.5-3B). After calculating slopes (\pm Standard Error) for each growth curve it became obvious that DOX addition has detrimental effect on cell growth rate; this is either due to killing cells, differentiation, or slower cell cycle. Moreover, PG1.1 SO c4 cells proliferated twice as fast as parental PG1.1 cells (slope of 2.26 versus 1.15). To control for these effects, in the following differentiation experiments we have used a lower seeding density to minimize spontaneous density induced differentiation.

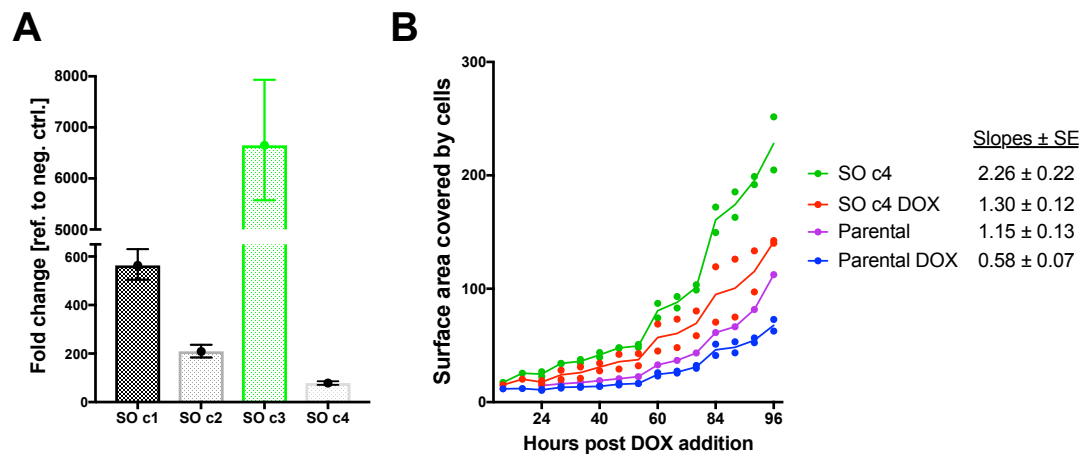


Figure 5-3 | Construction and characterization of PG1.1-SO clonal cell lines.

(A) qPCR data showing mRNA levels of *SOX10* in different PG1.1-SO clonal lines which have not been exposed to DOX, n=1. These are normalized to parental PG1.1 cell line (negative control) which has no *SOX10* expression (CT value 40) (B) Growth curves of PG1.1-SO_o c4 and parental PG1.1 cell lines in presence and absence of 1 μ g/ml DOX, n=2.

5.1.2 Construction of PG1.1 cell line for *Sox10* activation with dCas9-VP160 (PG1.1-3SON)

After constructing cell lines capable of *SOX10* and *OLIG2* overexpression (positive control) we went on to construct another cell line which contained sTFs targeting same OPC master regulators – idea being that we will compare these lines (and therefore cDNA and sTFs) head-to-head in their ability to drive PG1.1 differentiation.

To construct such line, we have transfected PG1.1 cells with 3SON (contains mRuby reporter gene and PiggyBac homology arms) and CAG_pBase. Cells were cultured for two weeks and sorted for mRuby (Fig. 5-4A). mRuby positive cells (now named PG1.1-3SON) were then further cultured as a pool of transfectants. Three weeks after initial transfection, mRNA was harvested and expression of *Sox10*, *Olig2* and *Nkx6-2* was tested using by qPCR (Fig. 5-4B). In comparison to PG1.1 parental line, *Sox10* mRNA levels in PG1.1-3SON cells were significantly upregulated as expected, levels of *Olig2* and *Nkx6-2* remained the same.

To summarize, we now have constructed (i) PG1.1-SO cell line which massively overexpress *SOX10* and *OLIG2* in response to DOX and has a small leakage of these transgenes in absence of DOX (20x lower in comparison to cells exposed to DOX); and (ii) PG1.1-3SON cell line which has sTFs targeting and activating endogenous *Sox10* expression (with no direct effect on other genes). Next, we compare these cell lines head-to-head in their ability to affect NSC differentiation.

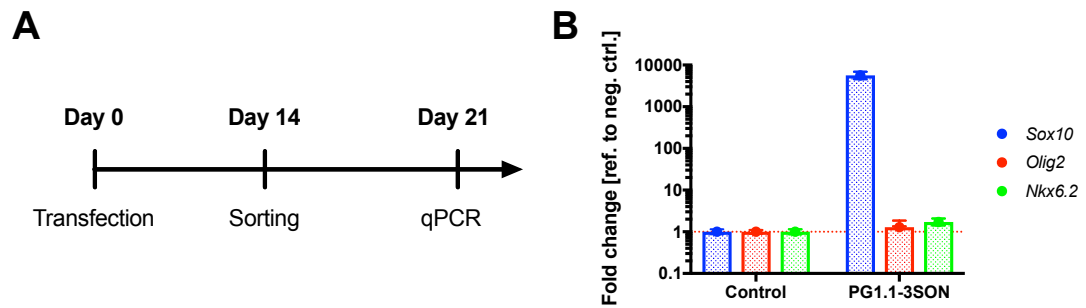


Figure 5-4 | Construction of stable PG1.1 cell lines containing 3SON sTFs (PG1.1-3SON).

(**A**) Graphical representation of the experiment. PG1.1 cells were transfected with plasmid 3SON and CAG_pBase and cultured for 2 weeks. mRuby positive cells (reporter present in 3SON plasmid) were sorted and cultured further. (**B**) qPCR data showing *Sox10*, *Olig2* and *Nkx6-2* mRNA levels in parental PG1.1 cells (Control) and Ruby sorted cells (PG1.1-3SON), n=1

5.1.3 Effect of *SOX10* and *OLIG2* cDNA overexpression and *Sox10* activation with sTFs on mouse neural stem cells in self-renewal conditions

We next tested whether the forced expression of *SOX10-OLIG2* and endogenous activation of *Sox10* can trigger the anticipated increases in OPC specification in NSCs. Following experiments were performed (Fig. 5-5A). PG1.1-SO c4 (*SOX10* and *OLIG2* cDNA overexpression) and PG1.1-3SON (endogenous *Sox10* transcription activation using sTFs) cells were seeded at very low density (Fig. 5-5B) into 6-well plates with normal NSC growth medium (containing EGF and FGF). Parental PG1.1 cells were used as a control. 24hr later, 1 μ g/ml of DOX was added to some of the PG1.1 and PG1.1-SO cells. These cells were then analysed by flow cytometry for presence of PDGFR α -GFP positive cells to determine whether these factors could override self-renewal conditions and trigger differentiation. At each time point, fresh self-renewing PG1.1 cells were used to set gating (for example, first box from left in Fig. 5-5D).

These experiments showed that as expected the parental PG1.1 cell line (Figure 5.5 blue) – which served as a negative control – showed no GFP positive cells at Day 2 or Day 4. By Day 5, we observed ~10-15% of cells becoming GFP positive, presumably due to the spontaneous differentiation triggered by confluence.

PG1.1-SO c4 cells (which expressed *SOX10* and *OLIG2* cDNA) increased GFP every day. At Day 2, PG1.1-SO c4 cells (with or without DOX) were on average around 7% GFP positive. This number increased over time and reached 24.2% to 47.1% at Day 5. Interestingly, we observed a difference in differentiation rates between PG1.1-SO c4 cells with (green) and without DOX (red). PG1.1-SO c4 cells

which have not been exposed to DOX (red) were able to differentiate to PDGFR α -GFP positive cells even at low confluence (Day 2 to Day 4). This could be explained by *SOX10* (but not *OLIG2*) expression leakage in these cells (Fig. 5-5C, red). Such findings suggest that *SOX10* alone is sufficient to drive NSC differentiation or at least prime the cells by turning on PDGFR α . Importantly, if DOX is added to these cells, the number of PDGFR α -GFP positive cells is lowered (Fig. 5-5C, green versus red).

We speculate that this could be explained either (i) by detrimental effect of extremely high levels of transgene expression or (ii) by high levels of *OLIG2* which might interfere with differentiation and push NSC to maintain their multipotency. Considering that *OLIG2* is one of the key master regulators of NSC identity, this could explain such observations. Thus, it seems that low level expression of *SOX10* cDNA is sufficient to direct NSC differentiation to OPC lineages.

Lastly, and most importantly, we tested if endogenous *Sox10* transcription activation by sTFs (dCas9-VP160) was also able to drive NSC differentiation towards OPC lineage in the PG1.1-3SON cell line (Fig. 5-5 D-E). In this cell line, *Sox10* transcription is constitutively activated by sTFs (pink in Fig.5-5C) to levels which are higher than *SOX10* cDNA overexpression without DOX (red in Fig.5-5C) but lower to those after DOX addition (green in Fig.5-5C). Importantly, such *Sox10* activation by sTFs was sufficient to drive NSC differentiation into PDGFR α positive cells (Fig.5-5 D-E).

The levels of forced NSC differentiation were similar between using *SOX10* cDNA (red) and sTFs (pink) targeting *Sox10* (46.9% and 50.8%, respectively), but significantly lowered if *OLIG2* overexpression (green) was included (30.2%).

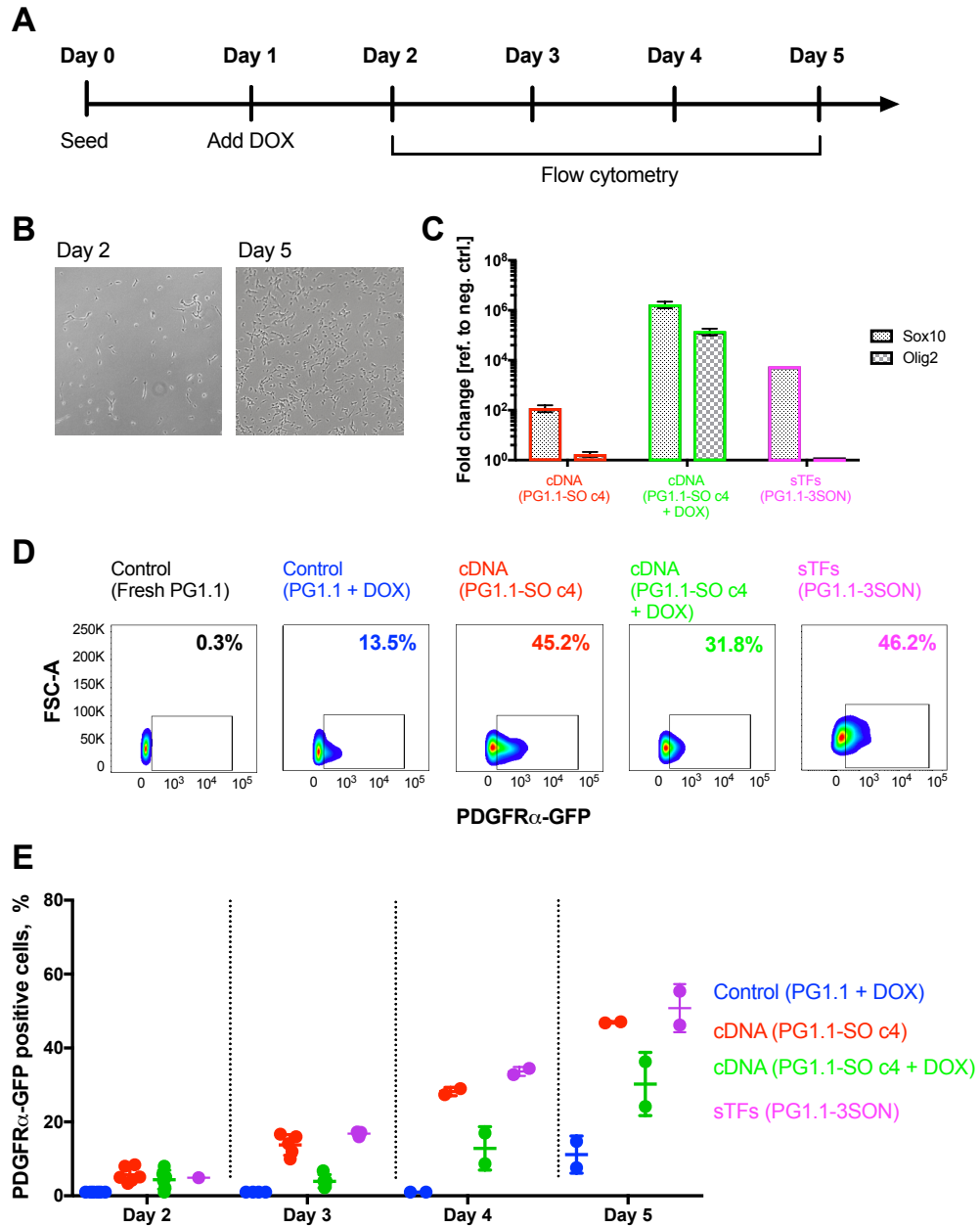


Figure 5-5 | Effect of *SOX10* and *OLIG2* overexpression and *Sox10* activation on mouse neural stem cells in self-renewal conditions.

(A) Schematic representation of the experiment. Cells were seed in 6-well plate at low density (2×10^4 cells/well). Next day, $1 \mu\text{g/ml}$ of DOX was added to some of control cells (PG1.1) and *SOX10-OLIG2* overexpressing cells (PG1.1-SO). Cells were harvested at 4 different time points. (B) Phase contrast images demonstrating cell density at Day 2 and Day 5. (C) qPCR data demonstrating levels of *Sox10* and *Olig2* mRNA at day 2, $n=1$. (D) Flow cytometry data showing amount of cells positive for PDGFR α -GFP at Day 5. Fresh PG1.1 cell (black) were used to set gating for positive cells. (E) Quantification of flow data, each dot represents biological replicate.

5.1.4 Effect of *SOX10* and *OLIG2* cDNA overexpression and *Sox10* transcription activation (sTFs) on mouse neural stem cells in OPC differentiation conditions

To test if *SOX10* overexpression and *Sox10* activation could also enhance NSC differentiation when delivered during conditions which favour differentiation, the following experiments were done (Fig. 5-6A). Cells were seeded at low density and induced to differentiated next day by addition of FGF, PDGF-AA and Forsklin. Some cells were exposed to 1 μ g/ml DOX. Four days after induction, cells were analyzed using flow cytometry for PDGFR α -GFP positive cells (Fig. 5-6B).

The following observations were made: parental cells (blue) showed around 20% of cells that became GFP positive as expected. PG1.1-SO c4 cells (*SOX10* cDNA overexpression; red) had on average 52.5% of cells GFP positive, while if these cells were exposed to DOX (*SOX10* and *OLIG2* cDNA overexpression; green), the number of GFP positive cells dropped to 31.4% on average. PG1.1-3SON cells (*Sox10* sTFs; purple) were able to generate over 70% PDGFR α -GFP positive cells on average (Fig. 5-6C).

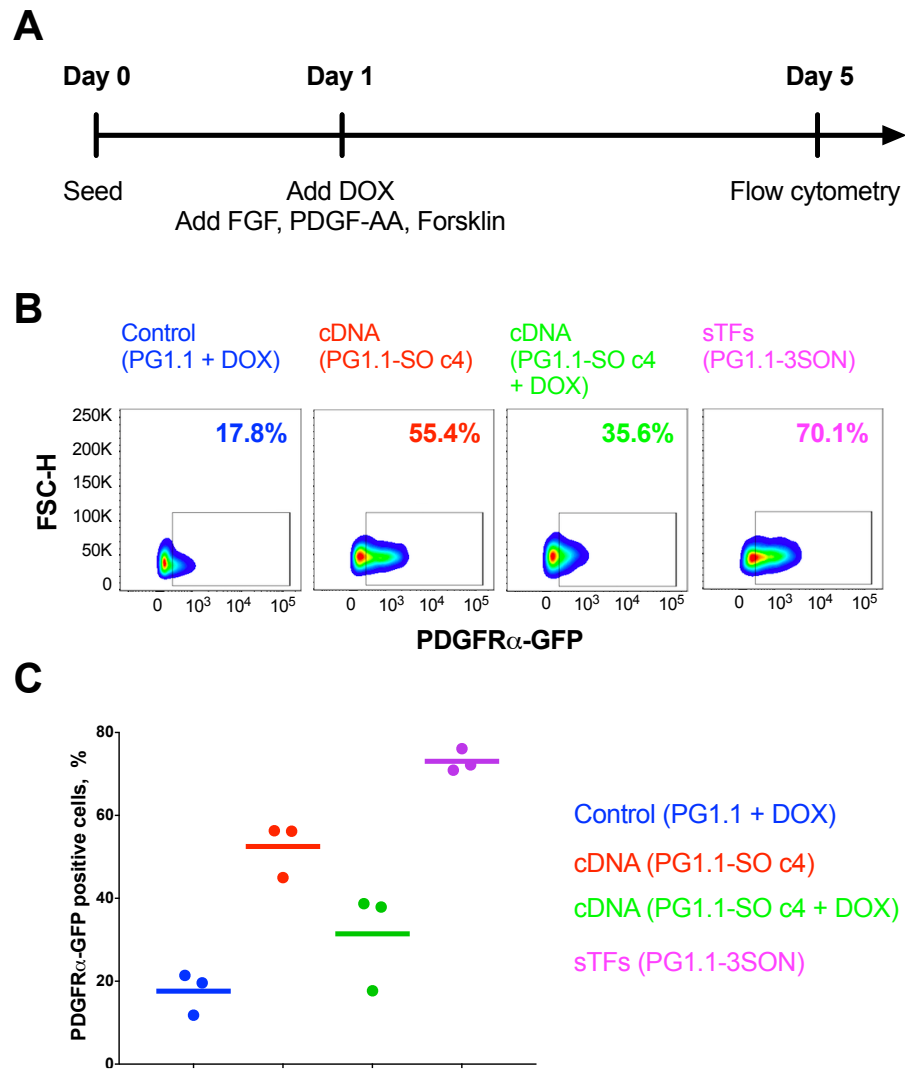


Figure 5-6 | Effect of *SOX10* and *OLIG2* cDNA overexpression and *Sox10* activation on mouse neural stem cells in OPC differentiating conditions.

(A) Schematic representation of the experiment. Cells were seed in 6-well plate at low density (2×10^4 cells/well). Next day, 1 μ g/ml of DOX and OPC differentiation medium (FGF, PDGF-AA, Forsklin) was added to cells. Cells were analyzed 4 days after differentiation induction. (B) Flow cytometry data showing amount of cells positive for PDGFR α -GFP at Day 5. (C) Quantification of flow data, each dot represents biological replicate.

Under these conditions cells which had *Sox10* transcription activated by sTFs (PG1.1-3SON) had more differentiating cells than cells with *SOX10* cDNA overexpression (PG1.1-SO c4 with no DOX). Such observation could be explained by higher *Sox10* levels in PG1.1-3SON cells compared to that of PG1.1-SO c4 without DOX (Fig. 5-5C, red and purple). Although the highest SOX10 levels were seen after PG1.1-SO c4 was exposed to DOX we might expect this condition to have most PDGFR α -GFP positive cells if differentiation is directly linked to SOX10 levels, but this is not a case. We speculate that in these cells (PG1.1-SO c4 with DOX) too high SOX10 and/or OLIG2 levels interfere with differentiation process. sTFs can therefore clearly program cell fate (NSC differentiation) in a comparable manner to cDNA overexpression.

To test if the GFP expressing cells are fully specified OPCs or are they just ‘primed’ NSC which have activated the marker PDGFR α we tested them for a number of additional OPC and OL specific markers: MBP, O4 and NG2. Parental PG1.1 (control) and PG1.1-SO c4 (*SOX10* cDNA overexpression) cells were seeded at low density (2×10^4 cells/6-well plate well) and induced to differentiate to OPCs by addition of FGF, PDGF-AA and Forskolin. Four days after induction, medium was changed to drive OL differentiation. This medium contained T3 and L-Ascorbic acid. Four days later, cells were analysed by immunocytochemistry (Fig. 5-7A). We have found that although PG1.1-SO c4 cells generated MBP and O4 positive cells, PG1.1 (parental) cells failed to do so. Moreover, morphology of O4 positive cells did not resemble standard OPC or OL morphology (Fig. 5-7B). For this reason, and to test observed effects of *Sox10* transcription activation in NSC with different genetic background, used an independent Black6 (BL6) NSC line.

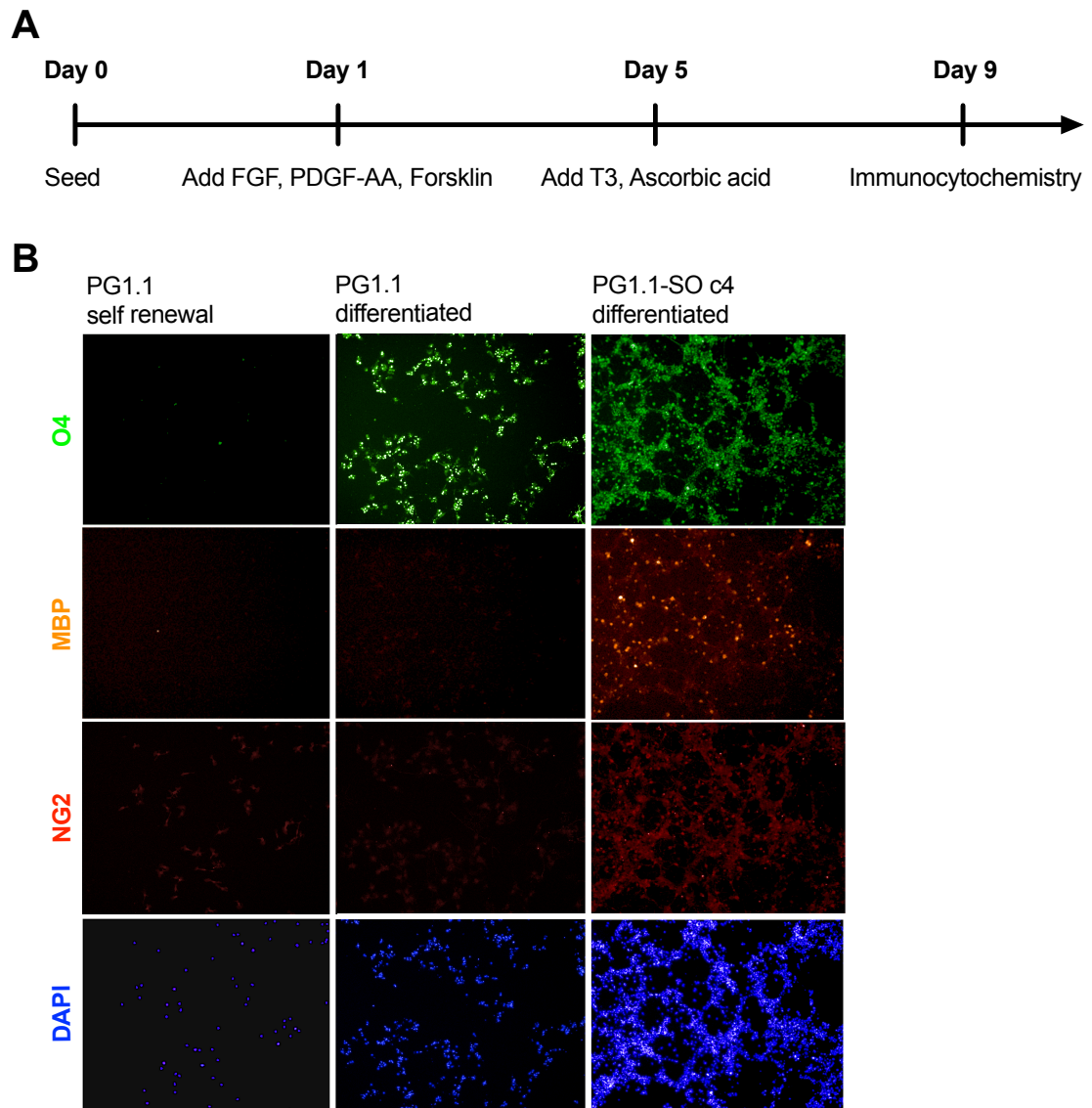


Figure 5-7 | Differentiation of PG1.1 and PG1.1 SO c4 NSC to OL lineage.

(A) Schematic representation of the experiment. (B) Immunocytochemistry of PG1.1 cells after differentiation. O4, MBP and NG2 antibodies have been used to identify successfully differentiated OLs.

5.1.5 Effect of *Sox10* transcription activation by sTFs on BL6 NSC differentiation towards OL lineage

To test observed effects of *Sox10* activation in NSC with an independent cell line with a different genetic background, we performed similar experiments using a newly established Black6 (BL6) NSC line. This is a non-transgenic wild-type primary adult NSC culture. BL6 cells were exposed to a range of different culture conditions known to induce differentiation, including: basal NSC growth medium lacking any growth factors, containing 2% FCS or standard OL differentiation conditions as described before. After ICC for O4, it was evident that standard OL differentiation conditions gave rise to large number positive cells, as well as some MBP positive cells (Fig. 5-8). The other two conditions gave rise to a lower number of OLs.

To repeat *Sox10* transcription activation experiments in BL6, we constructed a new sTF Ai1 plasmid (S10) which contained dCas9-VP160 and 8 gRNAs targeting *Sox10* promoter (Fig. 5-9A). By increasing the number of gRNAs targeting *Sox10*, we were hoping to increase the number of cells which have *Sox10* gene activated.

To make a stable cell line containing these sTFs targeting *Sox10*, we have transfected BL6 NSC cells with S10 and CAG_Pbase plasmids and cultured them for couple weeks. To test if *Sox10* could be activated in these cells, we harvested some RNA and performed qPCR. As expected, these cells had *Sox10* gene activated (Fig. 5-9B). Next, we sorted for mRuby positive cells (reporter for S10 plasmid) to enrich the population of cells with had *Sox10* activated. About 1.4% of cells were mRuby positive (Fig. 5-9C). This new cell line (BL6-S10) had normal NSC morphology (Fig.5-9D).

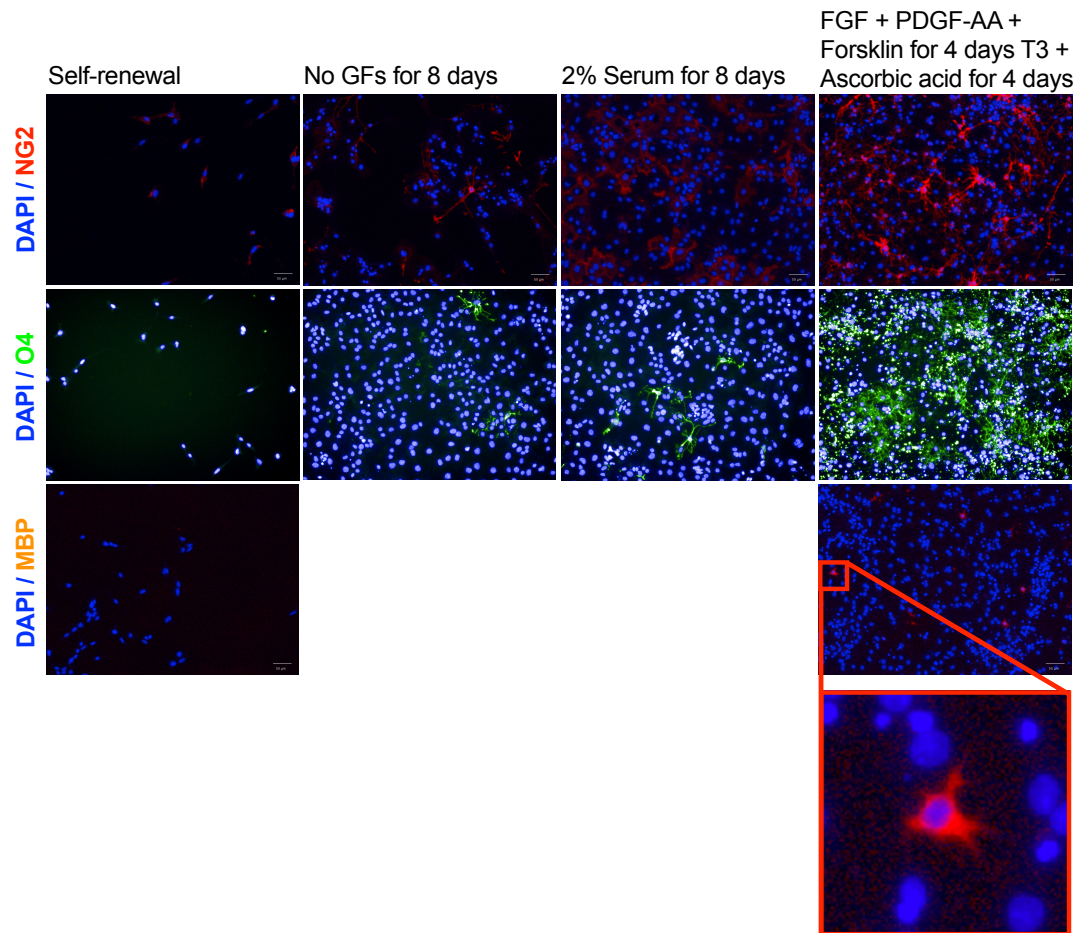


Figure 5-8 | Characterizing BL6 cell line differentiation potential in different differentiation inducing growth mediums.

Immunocytochemistry performed for NG2, O4 and MBP on BL6 cell 8 days after differentiation induction with one of the specified methods.

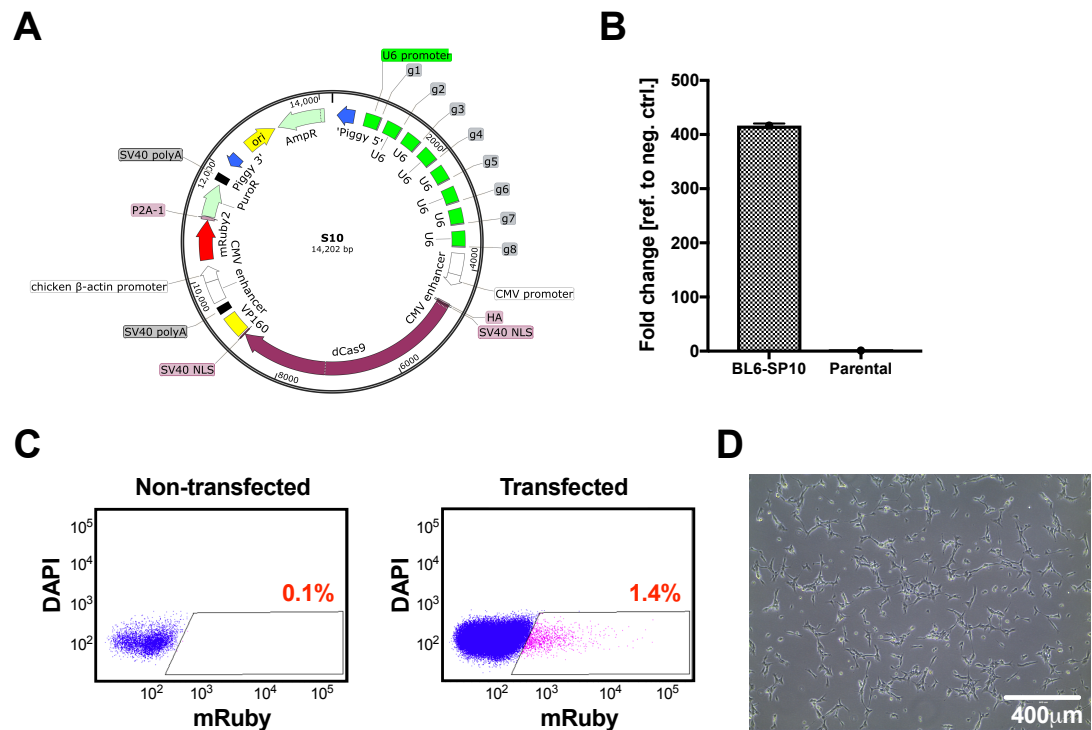


Figure 5-9 | Constructing stable BL6 cell lines expressing sTFs targeting *Sox10* (BL6-S10).

(A) Map for the S10 plasmid containing 8 gRNAs (g1-g8) targeting proximal *Sox10* promoter and dCas9-VP160. (B) qPCR data showing levels of *Sox10* mRNA in BL6-S10 stable cell line, n=1. (C) mRuby positive cells (reporter for S10 plasmid) were sorted 8 days after BL6 cell were transfected with S10 plasmid and CAG_Pbase. (D) Live phase contrast image showing normal BL6-S10 cell morphology.

While BL6 cells alone (parental/control) were able to generate O4 positive OPCs/OLs (Fig. 5-10B), we were interested to investigate if *Sox10* transcriptional activation with sTFs would be able to enhance this process. To test this, BL6 (control) and BL6-S10 (contain sTFs) cells were induced to differentiate to OPCs using standard 8 day two-step protocol used earlier. Cells were analysed for O4 positive cells using flow cytometry (Fig. 5-10C) and it was evident that there was no significant difference in a number of O4 positive cells between parental BL6 cells (blue) and BL6-S10 cells which had *Sox10* gene activated by sTFs (green). In both conditions, around 60% of cells were positive for O4 while BL6 in self renewal conditions had <1% of positive cells (Fig.5-10D).

As we could not observe any effect of the sTF-driven activation of *Sox10*, we reasoned that this differentiation protocol is already so efficient that system is saturated; *Sox10* activation cannot enhance it further. We therefore tested these two cell lines in sub-optimal differentiation medium (as shown in Fig. 5-8) and used 5% FCS, instead of T3 and L-ascorbic acid, to force last stage of differentiation (Fig. 5-10E). Parental BL6 cells (blue) only gave 10.1% of O4 positive cells on average, while BL6-S10 cells (contain sTFs) gave 27.4% O4 positive cells on average (Fig. 5-10F). Thus, *Sox10* activation by dCas9-sTFs has a clear effect on NSC differentiation towards OPC and OL lineage, but this effect is only significant when using sub-optimal OPC differentiation conditions.

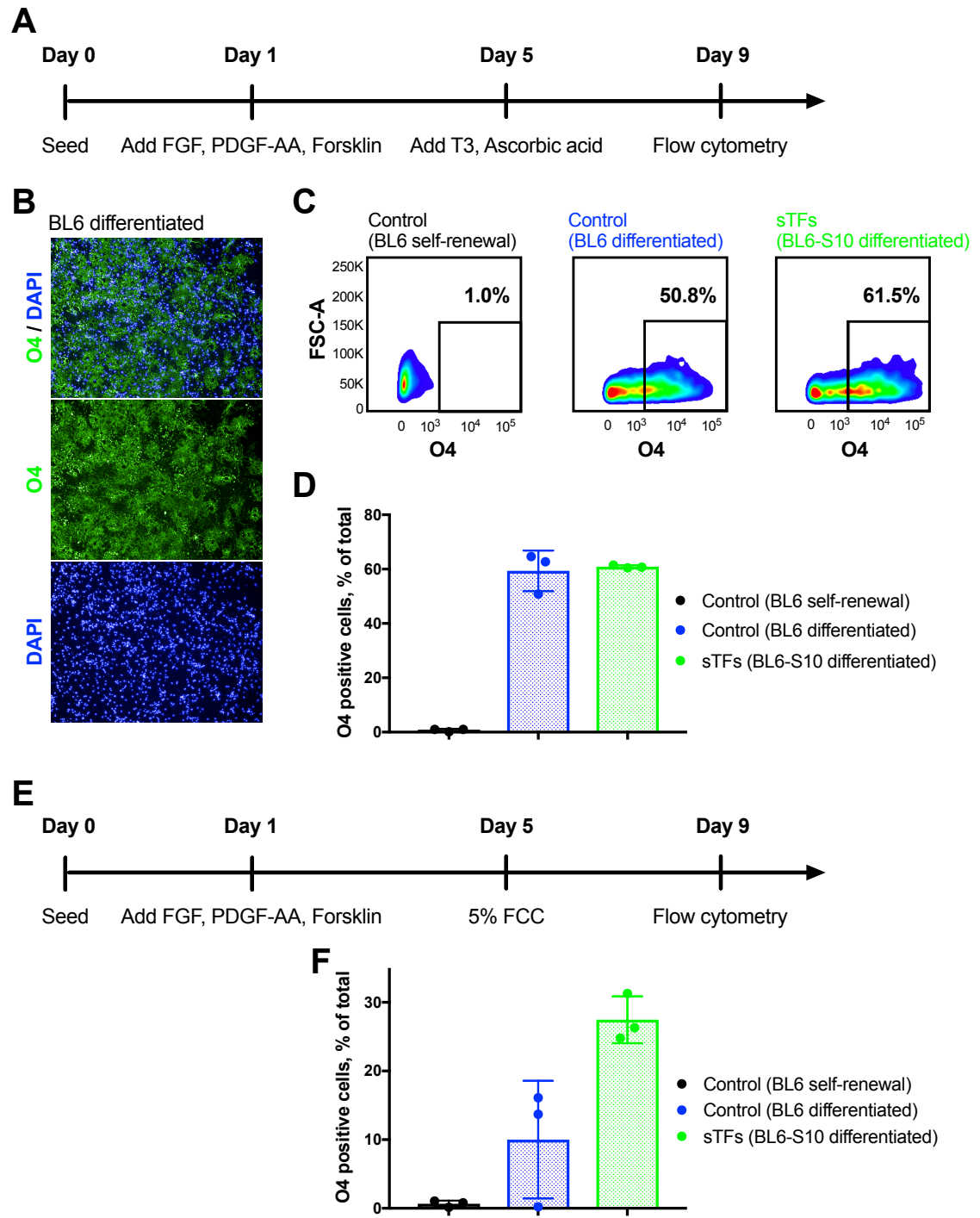


Figure 5-10 | Effect of *Sox10* activation on NSC differentiation to OL lineage.

(A) Schematic representation of first set experiments. (B) Immunocytochemistry images after O4 staining on differentiated BL6 NSC cells at Day 9. (C) Flow cytometry data of BL6 and BL6-S10 cells after the differentiation. (D) Quantification of flow cytometry data of experiment described in A, each dot represents biological replicate. (E) Schematic representation of second set of experiments. (F) Quantification of flow cytometry quantification of experiment described in E, each dot represents biological replicate.

5.2 Effect of *Sox10*, *Olig2* and *Nkx6-2* transcription activation on MEF transdifferentiation to OL

5.2.1 Generation of MEF lines from mouse embryos

Given the sTFs could affect differentiation program in NS cells, we next tested if they could trigger the more challenging reprogramming event: transdifferentiation of MEFs to OPCs. For the following transdifferentiation experiments, numerous fresh MEF cultures were generated from 13.5-day old mouse embryos (Fig. 5-11A), by dissecting distal limb fibroblasts (Fig. 5-11B). The full protocol can be found in Method and Materials chapter.

MEFs showed standard fibroblast morphology (Fig. 5-11C). We have compared levels of *Nkx2-2*, *Nkx6-2*, *Sox10*, *Olig2*, *Ascl1* and *Sox9* in two separately generated MEF lines to those in NSC (Fig. 5-11D). All cells had low levels of *Nkx2-2*, *Nkx6-2* and *Sox10*. MEFs also had low levels of *Olig2* and *Ascl1*. All cells had high *Sox9* levels.

MEFs are known to be hard to transfect using Nucleofection. Therefore, we have tested transfection efficiency with different reagents. Transfection with PEI showed only 10% transfection efficiency (data not shown), while transfecting CAG_GFP to MEFs with Lipofectamine 3000 gave almost 30% of transfection efficiency as demonstrated by flow cytometry data (Fig. 5-11E). This method was used in subsequent experiments.

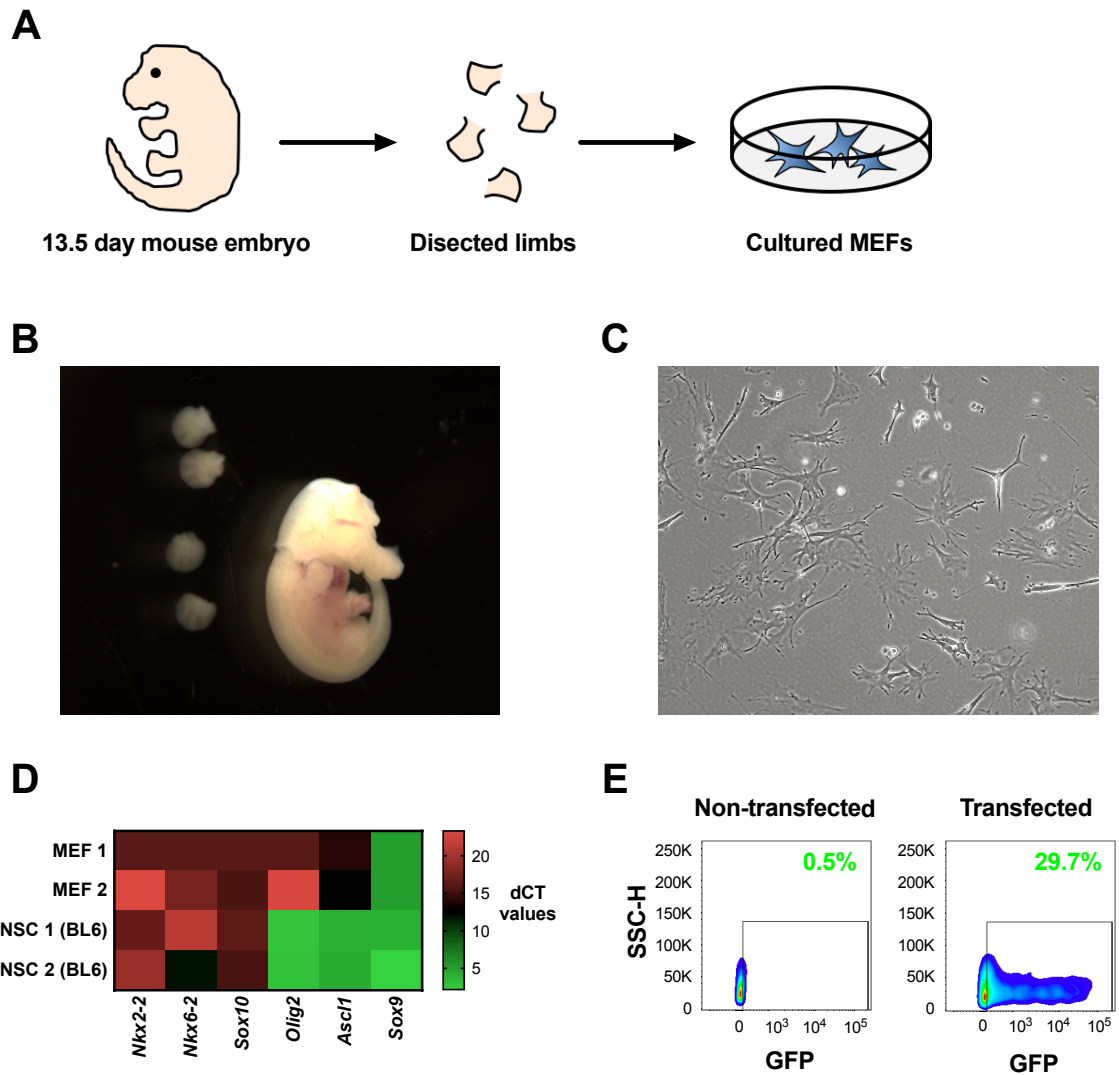


Figure 5-11 | Generating MEF lines from mouse embryos.

(A) Schematic representation of MEF harvesting method. Limbs are removed from the 12.5-day old mouse embryo and placed in the 0.25% Tripsin/EDTA for dissociation. Cells then are plated and cultured in AmnioMax medium. (B) Image showing limb separation from the embryo. (C) Phase contrast image of cultured MEFs 10 days after harvesting. (D) dCT values for various genes in MEF and NSC, n=1. (E) Flow cytometry data of non-transfected MEFs and MEFs transfected with CAG_GFP reporter plasmid three days before analysis using Lipofectamine 3000 kit, n=1.

5.2.2 MEF transdifferentiation to OPC/OLs after *Sox10*, *Olig2* and *Nkx6-2* activation with sTFs (3SON)

To test if *Sox10*, *Olig2* and *Nkx6-2* activation can force MEF transdifferentiation, the following experiments have been performed (Fig. 5-12A). MEFs were seeded and next day either transfected with 3SON and CAG_Pbase or not transfected. Three days after transfection, some cells were harvested for analysis by qPCR, while the remaining cells were incubated in MEF transdifferentiation medium (see Methods and Materials). Some cells were also harvested for qPCR at Day 9. At day 14, some cells were harvested for qPCR while remaining cells were induced to differentiate to OLs for 3 days. Subsequently, cells were analysed by flow cytometry and immunocytochemistry at Day 17.

From the qPCR marker analysis, it was clear that activation of all three genes was occurring at each of the three time points; thus, transfection and target gene activation by the sTFs was successful (Fig. 5-12B). Phase-contrast images, at Day 14 indicated that transfected cells had morphological changes from flat MEF-like shape to spindle OPC-like shape in a small subset of cells (Fig. 5-12C).

To test if MEFs induced to transdifferentiate expressed the OPC/OL marker O4, we have conducted ICC and flow cytometry at Day 17. ICC indicated only very small number of cells were O4 positive; yet, encouragingly, these cells did display a morphology reminiscent of OPCs (Fig. 5-13A). After analysing transfected and non-transfected MEFs, we found that transfected MEFs had 3.2% of O4 positive cells, while non-transfected MEFs only had 0.9%. Although there is a small difference

between two conditions, this is too small to be confident that transdifferentiation happened.

We hypothesised that levels of gene transcription might have been too low for efficient transdifferentiation. In an attempt to boost transcription levels, we tested additional combinations of sTFs and cDNA. In addition to using 3SON sTFs, we have also included three other conditions: overexpression of *SOX10* and *OLIG2* cDNA by itself; 3SON sTFs supplemented with overexpression of *SOX10* and *OLIG2* cDNA; and 3SON sTFs together with the addition *Sox10* sTFs (S10 plasmid) – to boost levels of *SOX10*. Secondly, we increased transdifferentiation times from 17 days to 24 days (Fig. 5-14 A).

After repeating this transdifferentiation, we indeed have found cells which morphologically are similar to OPC/OL cells in all conditions, except in non-transfected but induced control (Fig. 5-14 B). More importantly, after analysing these cells by flow, we have found that all conditions had over >10% of O4-positive cells, while less than 1% O4-positive cells were found in controls (Fig. 5-15C). Also, by increasing transdifferentiation protocol from 17 to 24 days, we manage to double the amount of OPCs generated when using 3SON sTFs (3.2% to 6.9%). Moreover, by adding additional sTFs targeting *Sox10* (and thus increasing *Sox10* expression levels), we managed to increase efficiency of transdifferentiation process from 6.9% to 17.9%. Thus, sTFs expression in MEFs can be sufficient to trigger transdifferentiation to OPCs.

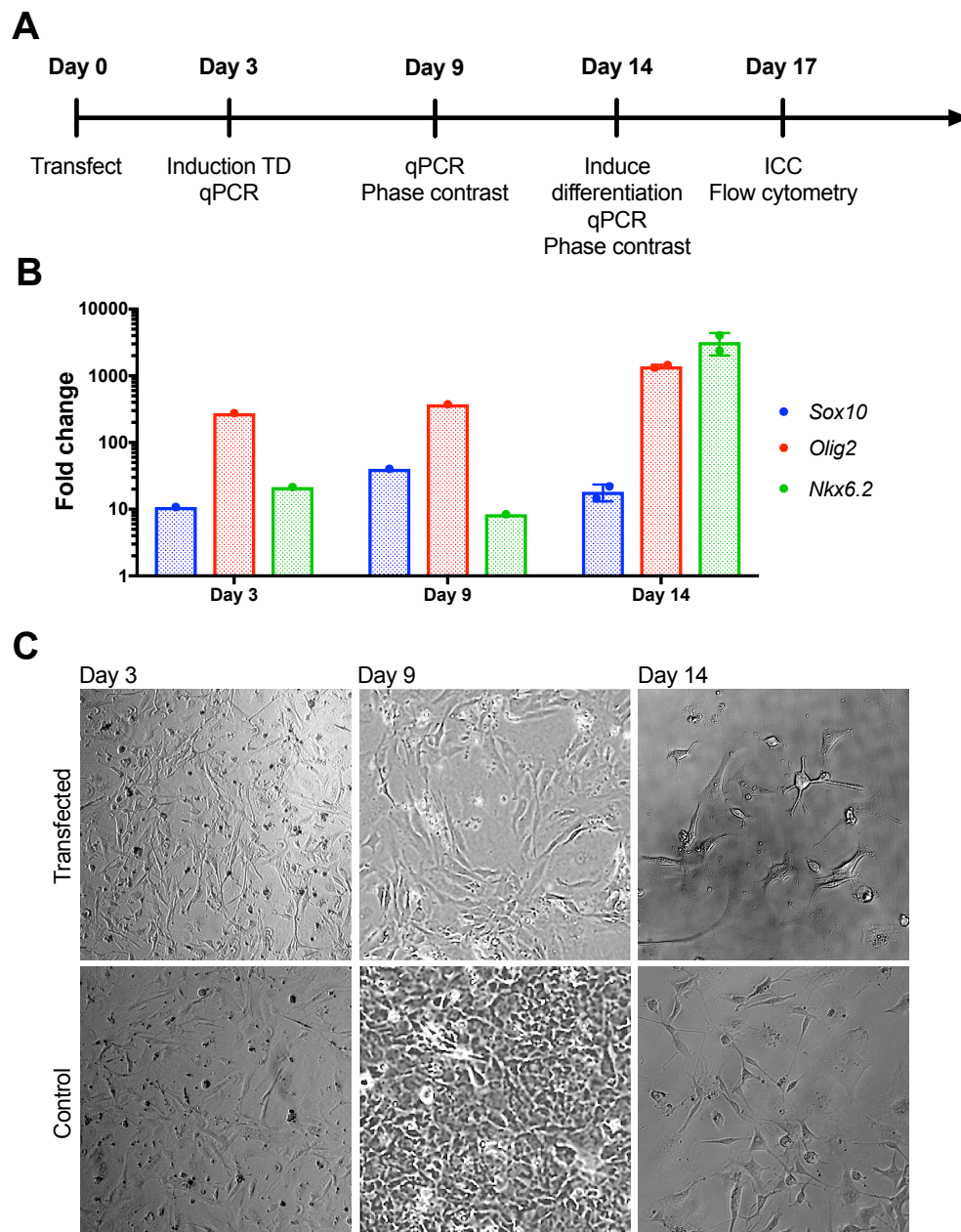


Figure 5-12 | MEF transdifferentiation to OPCs/OLs after *Sox10*, *Olig2* and *Nkx6-2* activation with 3SON.

(A) Schematic representation of the experiment. Cells are transfected on Day 0 with 3SON vector and CAG_pBase. At Day 3, some cells are harvested for qPCR, while the rest is induced to transdifferentiate by addition of MEF TD medium. At Day 9 and Day 14, some cells are analysed by qPCR. Transdifferentiation events are scored using immunocytochemistry and flow cytometry at Day 17. (B) qPCR data showing levels of *Sox10*, *Olig2* and *Nkx6-2* mRNA 3, 9 and 14 days after transfection, each dot represents biological replicate. (C) Phase contrast images at Day 3, Day 9 and 14 of Control and Transfected MEFs.

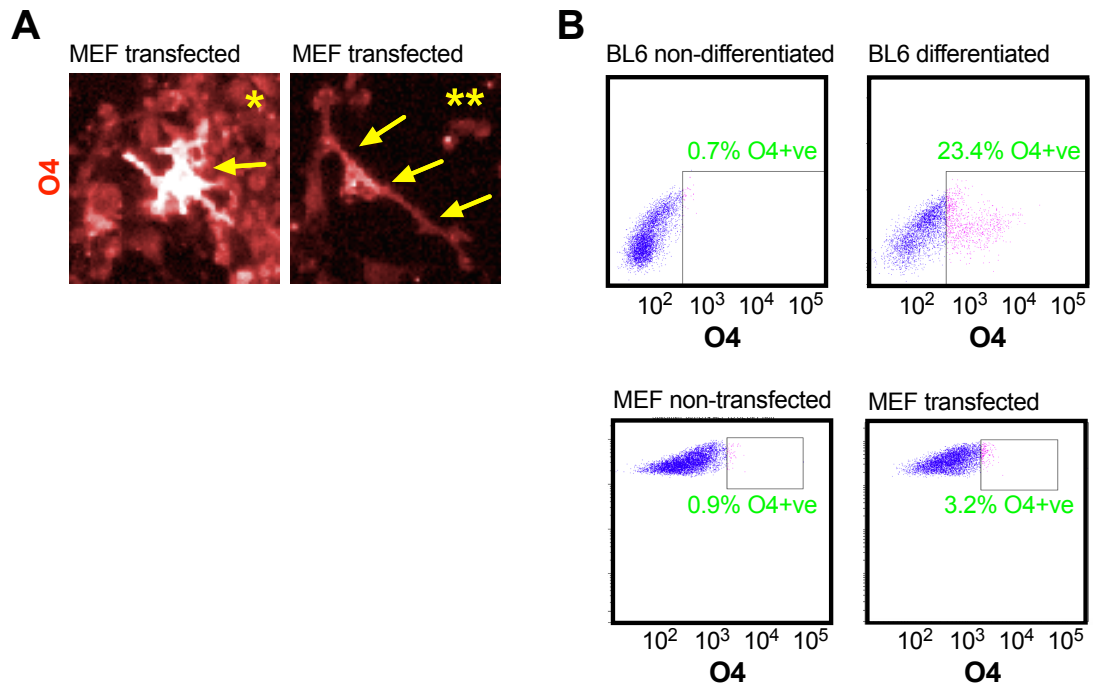


Figure 5-13 | MEF transdifferentiation to OL after *Sox10*, *Olig2* and *Nkx6-2* activation with 3SON.

(A) Immunocytochemistry of O4 in transfected MEFs at Day 17. (B) Flow cytometry for O4 in transfected and non-transfected MEFs 17 days post transfection. For positive control, differentiated BL6 cells has been used.

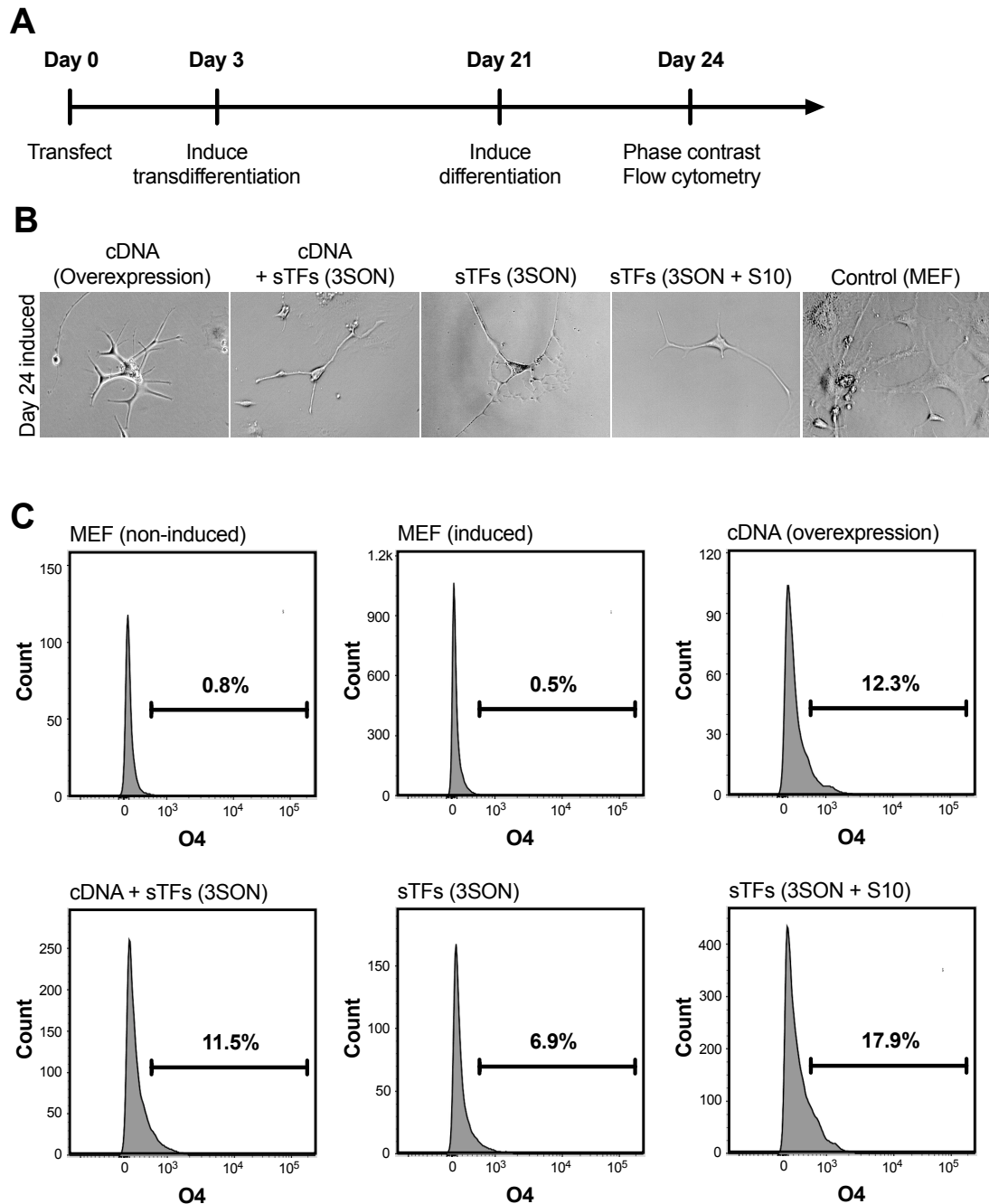


Figure 5-14 | Optimized MEF transdifferentiation to OL after *Sox10*, *Olig2* and *Nkx6-2* activation with sTFs.

(A) Outline of experimental workflow. MEFs have been transfected with following plasmids: “cDNA (Overexpression)” (CAG_Sox10_Olig2 plasmid); “cDNA + sTFs (3SON)” (CAG_Sox10_Olig2, 3SON and CAG_PiggyBac plasmids); “sTFs (3SON)” (3SON and CAG_PiggyBac plasmids); “sTFs (3SON + S10)” (3SON, S10 and CAG_PiggyBac plasmids). (B) Phase contrast images of OPC-like cells in different conditions at Day 22. (C) Flow cytometry data for levels of O4-positive cells in different conditions, day 24, n=1.

5.3 Concluding remarks

In this chapter we have explored if the sTFs produced in Chapter 3 and 4, have biological activity in redirecting cell fate – either by programming NS cell differentiation, or directly reprogramming MEFs to OPCs.

We successfully confirmed that overexpression (cDNA) or activation (sTFs) of *Sox10* in mouse NSC drives them towards OL lineage. In two different cell lines, we have increased both PDGFR α and O4 positive cells after SOX10 overexpression or activation of *Sox10* endogenous gene. Importantly, this effect was observed both in self renewal conditions and in differentiation inducing conditions. However, we have also observed that if differentiation level is high due to optimal differentiation conditions, *Sox10* activation does not have any observable effects on the number of O4 positive cells appearing at the end of differentiation process.

Secondly, we have demonstrated a successful transdifferentiation of MEFs into OPC-like cells, which was the main aim of this research. We have observed O4-positive cells arising from transdifferentiating MEFs after activating *Sox10*, *Olig2* and *Nkx6-2* with sTFs. At 17 days post transfection, we have found 3.2% O4-positive cells while at day 24 we have found 6.9-17.9% of O4-positive cells depending of what sTFs have been used. Furthermore, these cells had OPC-like morphology and were distinguishable from MEFs. Remarkably, these sTFs can activate target genes in a manner that has profound consequences on cellular phenotypes.

6 General discussion

6.1 Understanding dCas9 synthetic transcription factors

CRISPR/Cas9-based sTFs have been under intensive development in the past few years. To date, they have been used in proof of principle studies across a number of applications and our understanding of how they work is increasing rapidly. In this thesis, we have explored their utility in reprogramming and have identified both opportunities and limitations of current sTF approaches. We report successful activation of multiple master regulators in both MEFs and NS cells as well as an efficient method for construction of all-in one system for multiplexing. Importantly, we demonstrate that sTFs can direct both differentiation and transdifferentiation of cells.

6.1.1 sTFs have cell-type specific functional constraints that limits predictive design rules

It is known that a number of biological constraints limit sTF activity. In our studies, we have confirmed and extended past studies. We noted a large variation in activation potential between biological replicates, variation between cell types, and an effect of basal transcription level of the target gene as key factors influencing sTF activity.

Firstly, it became evident from early experiments and was observed throughout all our experiments that there is a large variation of how well we can activate the gene from one experiment to another. In some cases, variable responses could be explained by technical limitations. For example, it is hard to compare activation levels at population level in MEFs and NSC as transfection efficiency differs greatly between

these two cell lines (>20% and >70%, respectively). However, even when controlling for this and many other conditions (e.g. transfection efficiency, qPCR variability and cell density) we still observed very different levels of activation; there seems to be inherent variability in the ability and extent to which cells activate target gene transcription. This may vary as much as couple hundred-fold difference between seemingly identical conditions. We would speculate a major contributor to the variability is the inherent stochastic nature of transcription and sTFs as well as transcription bursting. Varying levels of sTFs on a per cell basis are likely to dramatically influence the level of activation and is difficult to control using plasmid transient transfection.

We have observed significant differences in ability of same sTFs to act in different cell types. An example of this variability is *Nkx6-2*. This is not expressed in either in MEFs or NSC. However, while we were able to achieve significant transcription activation in MEFs, we could not observe any activation in NSC. In addition, *Nkx2-2* gene activation shown similar patterns. Furthermore, activation of *Sox10* was variable when assessed between different NS cell lines: BL6 NSCs showed a weaker activation compared to PG1.1 NSC. Thus, the cell context and likely transcriptional and epigenetic status of the host cell drives the major differences in response to the same sTF. The major challenge to making effective sTFs lies in understanding the competence of the host cell. As 3D topology, enhancer usage, promoter usage, TSS, pre-existing transcriptional circuits, and epigenetic marks (DNA methylation, chromatin state) will all influence gene expression – it remains unclear which, if any, of these is the dominant predictor of sTF activity.

Chromatin accessibility and epigenetic modifications, including DNA methylation and nucleosome positioning, might be one of the most important factors. Compacted chromatin has been shown to prevent transcription factor binding in a number of cases (Iwafuchi-Doi & Zaret 2016). That being said, a number of studies have shown that CRISPR/Cas9 is able to bind condensed chromatin (Polstein et al. 2015). It is still possible that although Cas9 is able to bind condensed chromatin, it might not be able to function as sTF in such environment – competence of the TAD to recruit Pol II preinitiation complex might be impaired. Another important epigenetic factor to consider is nucleosome positioning. It has been reported by a number of studies that nucleosomes strongly interfere with CRISPR/Cas9 activity (Hinz et al. 2015; Horlbeck et al. 2016; Isaac et al. 2016). In the study by Horlbeck and colleagues, they have demonstrated that in vitro, not only Cas9 cannot cut on DNA sites containing nucleosomes but that even binding of dCas9 is inhibited by nucleosomes. Importantly, after addition of nucleosome remodelling enzyme *yChd1*, ability of Cas9 to bind and cut DNA was restored (Horlbeck et al. 2015). As nucleosome positioning might differ between cell types, this could explain some of the biological constraints observed. It would be interesting to tether *yChd1* enzyme to dCas9 as an additional effector and see if this would give dCas9 a stronger pioneering function. Thirdly, it might be possible that genes exploit different promoters and/or TSS sites in different cell lines. That might explain why *Nkx6-2* can be activated in MEFs but not in NSC. Fourthly, the presence of other regulators actively interacting with promoter might also play an important role; dominant repressive factors (transcription factors, repressive miRNAs etc.) may be expressed in NSC, but not MEFs, to regulate *Nkx6-2*, this might interfere with functional sTFs. If this is true, using repressive sTFs to repress the repressors

might be a solution. Moreover, key activating co-factors might be missing that are expressed in cell type specific manner. This could result in transcriptional pausing and therefore inability of sTF to fully activate target gene transcription. Lastly, mRNA stability and translation capacity might be important factors to consider. Inability to maintain transcribed mRNA and prevent its decay might explain why we cannot detect any mRNA increase after targeting promoters. Further works needs to be performed to test such hypothesis.

One consistent observation from our studies – and also observed in past work (Konermann et al. 2014) – is that a gene that already being actively transcribed, cannot be easily activated further. A clear example of this is activation of *Olig2*. *Olig2* gene is silent in MEFs but highly expressed in NSC. We could significantly activate transcription of *Olig2* in MEFs using sTFs; however, the same sTFs had no effect on transcription in NSC, as levels were already very high. To investigate this further, we decided to lower *Olig2* transcription in NSC by forcing them to differentiate after delivering sTFs. Some interesting observations were made. In 3 out of 4 experiments we saw that differentiating NSCs (containing sTFs) had higher percentage of OLIG2 positive cells as compared to differentiation NSC lacking sTFs. This would suggest that *Olig2* activation was successful. However, on average there was no statistical difference between these two conditions. One of the reasons for such contradictory observations' is the variability in the differentiation process itself (typically 5-20%). Thus, we do not have enough evidence to claim that *Olig2* activation in NSC during differentiation was reliable and significant.

The main mode of action for sTFs to activate transcription is by recruiting hosts' transcription machinery to the promoter. Therefore, it is not surprising that sTFs

have no effect on genes which already have such machinery recruited and operating. sTFs maybe therefore be most productive in recruiting the basal machinery to activate a non-expressed gene; but they are not able to increase levels of expression. It might be possible that sTFs would still have an effect in cases where transcription is happening but the rate is limited because some of the transcription partners are missing or restricted. In such cases, sTFs might compensate for lacking parts and increase the rate of transcription. Moreover, it might be possible that novel sTFs which act to remove transcription rate constraints would be able to force gene over-activation, but it remains to be seen if this is possible. Thus, our initial experiments showed sTFs can work extremely effectively, but mainly for non-expressed genes (with dramatic influence of host cell on their effectiveness); already expressed genes cannot be activated to higher levels.

6.1.2 Delivery of multiple gRNAs increases the success of transcriptional activation

In all cases we have observed the use of multiple gRNAs, which target very proximal sites on the promoter, gives better gene activation as compared to using a single gRNA. The first examples of this came from using pools of 10 different gRNAs to target either *Sox10*, *Olig2* or *Nkx6-2*. These pools strongly outperformed individual gRNAs as measured by qPCR. Furthermore, combining 2 to 3 gRNAs together also showed greater fold change compared to those gRNAs individually in all instances that we have tested. This suggests that some kind of cooperative effect exist between different gRNAs. However, as all of these observations have been made by co-

transfecting relatively large numbers of plasmids it is hard to control for homogenous plasmid delivery. Moreover, the assay of activation is performed at population level, providing only the average level across cells.

Do multiple gRNAs activate genes to higher level or do they just activate the gene in more cells? To answer this we performed single cell analysis on *OLIG2* activation in HEK293T cells. In these experiments, we used plasmids which contained either 1, 2 or 3 different gRNAs. Similar to past observations, at population level the more gRNAs are delivered, the higher fold-change is observed. Moreover, fold-change differences could be consistent with a synergistic effect of gRNAs (individual gRNAs activate ± 30 -fold while all three gRNAs together activate to ± 300 -fold). However, when we explored levels of OLIG2 protein in individual HEK293TT cells, we do not see synergetic patterns. Although, we see more OLIG2 positive cells using multiple gRNAs (52.9%), this was not dramatically different from the single gRNA ($\pm 40\%$ of OLIG2 positive cells). Therefore, the effect of multiple gRNAs seems to be an increase in a number of individual cells being activated. A similar trend is observed after measuring OLIG2 intensity in OLIG2 positive cells: enhancement/cooperation is observed if multiple gRNAs are used, but the difference is not significant enough to suggest synergy. To sum up, although at population level (qPCR data) synergy seems to exist, this is not confirmed by single-cell protein data. From single cell protein level analysis, we found that using multiple gRNAs increases both the number of cells that are activating the gene and the level of activation in those cells. However, although statistically significant, increase is too small to suggest synergy. Rather, a cooperation between different gRNAs to slightly increase activation potential is a better explanation. At population level, this might create an illusion of synergy.

There are some unknowns in these experiments which we could not control technically and should be considered. Although we co-delivered gRNAs as a single vector, we do not know the absolute levels of expression for each gRNA in each cell. Furthermore, the assembly of the dCas9-VP160/gRNA complex is another variable which we cannot technically control or measure. However, we would not expect either of these to differ significantly.

Based on the observations we have made, we propose the following working hypothesis on how multiple gRNAs cooperate at the transcriptional level. Firstly, by increasing the amount of gRNA, the chance of dCas9-VP160/gRNA complex formation (Fig.6-1A-B, blue arrow) or complex binding to the DNA is increased (Fig.6-1 A-B, pink arrow). At some point, this reaches threshold; increasing gRNA amount further has no further effect. However, by introducing new gRNA sequences which binds to surrounding DNA, the chances of complex forming and binding to DNA increases even further (Fig.6-1C). Thus, a diverse set of gRNAs is likely to have a greater chance of accessing and binding in the dynamic region around the promoter; i.e. there is greater chance of binding in a permissive state. Having multiple targets on the promoter allows dCas9-VP160/gRNA complex to bind to the promoter even if some of the target sequences are temporary restricted (Fig.6-1D), which would strongly interfere with sTF activity if only single target (gRNA) is available. Lastly, to initiate transcription, sTF has to attract multiple partners to the TSS. By increasing the number of sTFs bound on the DNA, the chance and the rate of transcription machinery assembly might increase as well.

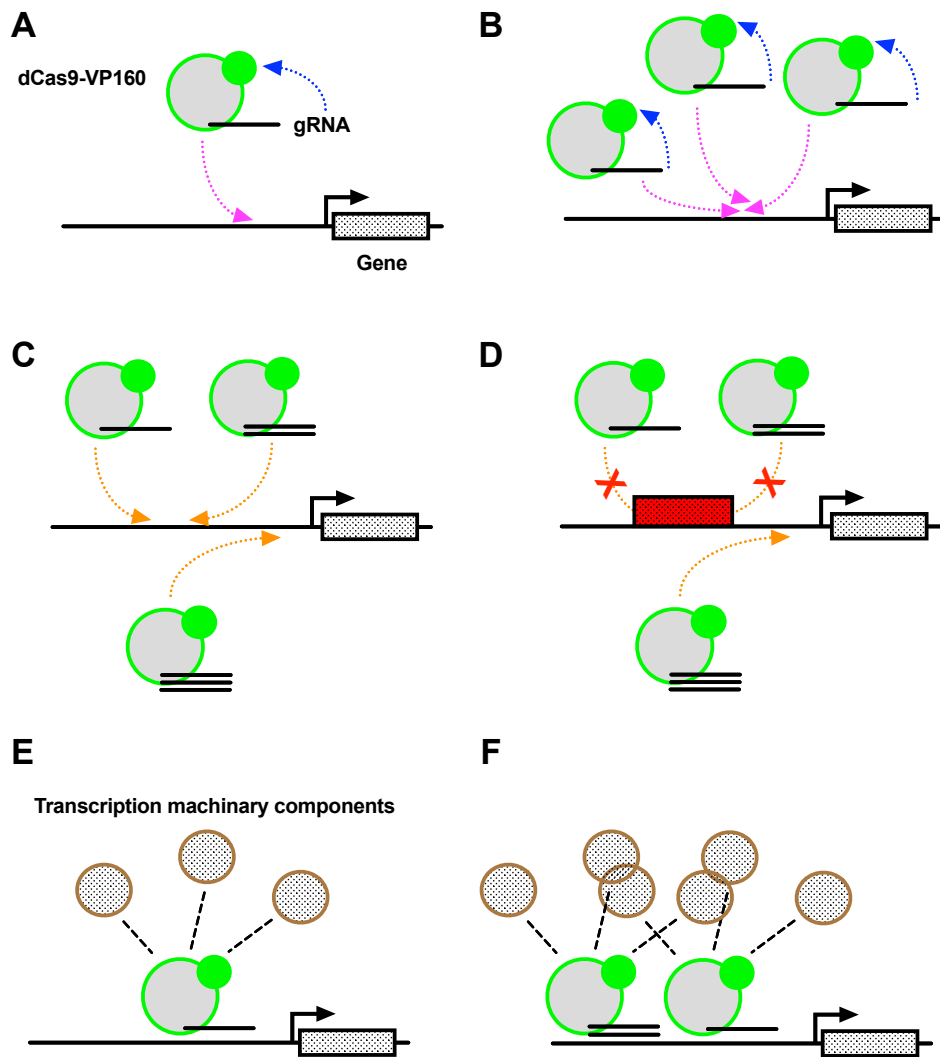


Figure 6-1 | Working hypothesis for how multiple gRNAs enhance transcription activation.

(A) A single sTF has to bind DNA to facilitate transcription initiation. At this stage, dCas9-VP160/gRNA complex formation (blue arrow) and complex binding to/dissociation from DNA (purple arrow) are some of the function defining reactions. (B) By increasing the amount of gRNA and dCas9, both of the reactions can be enhanced therefore enhancing activation potential. (C) By adding new gRNAs to the system and increasing targetable area, another rate enhancing feature is created. (D) Moreover, by providing additional targets when using multiple gRNAs, the chances that all of the targeted DNA sites will be blocked is decreased therefore increasing the chances of strong activation. (E) For a sTF to operate it needs to recruit a number of co-factors (like RNA PolII, TFIID etc.) to the promoter. (F) By increasing a number of sTFs, a chance to attract all required factors is increased.

6.1.3 dCas9-VP160 amount does not seem to be a key limiting factor

We have looked for a correlation between amount of dCas9-VP160 and the amount of target gene (OLIG2) after the activation in HEK293T cells. We observed low positive correlation. Certain cells which had high levels of dCas9-VP160 had low levels of OLIG2. In other cells, opposite was observed. This was further confirmed by RNA-FISH data in MEFs showing that some cells which have high amounts of *Sox10* mRNA have low amounts of dCas9 mRNA and vice versa. Thus, while the diversity and the absolute amounts of gRNAs plays a critical role in transcription activation, the absolute amount of dCas9-VP160 does not seem to be a limiting factor.

These data suggest that although presence of dCas9-VP160 is essential for transcription activation, its' levels are not limiting for the target transcription activation. Considering that only a single dCas9-VP160 copies are needed at the target site, and that high levels are produced from the expression plasmids, it is easy to understand why dCas9 levels are not limiting factor in gene activation. Our findings further suggest the diversity of gRNAs used is more important.

6.1.4 Stronger TADs allow higher gene activation but are still restricted by biological barriers

Another important consideration in design of effective sTFs is the effector that is attached to the dCas9 protein. We mainly used VP16 or multi-copy variant of it (VP160). Recently, a number of different novel TADs have been reported, which show significantly stronger activation of targeted genes than VP16 variants. We tested a number of different TADs in our systems side-by-side. We observed that TADs like

p300 and VPR are much stronger activators of some genes. In particular, p300 and VPR activated *Sox10* in NSCs to significantly higher levels compared to VP160. One of the likely reasons why VPR and p300 outperforms VP160 is by increasing the variation and the number of effectors which can attract transcription machinery, sTF is made more potent transcription activator. This is the same reason why VP16 outperforms VP160. In other words, this effect is similar to that of using single gRNA versus multiple gRNAs – the chances of recruiting the necessary basal transcription machinery is increased. Other groups have reported that increasing the amount and variation of TADs beyond that what VPR has no effect on activation potential. Such observations suggest that TAD potency is no longer a key limiting factor after certain point.

Although using stronger TADs works really well in some cases, at other times it has only minimal effects. One of the examples is *Nkx2-2* activation in MEFs. *Nkx2-2* is not transcribed in MEFs and it is possible to activate it with VP160. However, level of activation is relatively low. Interestingly, if VPR is used instead of VP160, activation is not enhanced as significantly as we have seen with *Sox10* activation in NSC. This suggests that in some cases biological constraints remain dominant, even for the stronger TADs. Another example which also supports this finding is *Nkx2-2* and *Nkx6-2* activation in NSC. Although both genes can be activated in MEFs even with VP160, neither of the genes could be activated in NSC either with VP160, p300 or VPR. Stronger TADs therefore can be useful, but do not consistently enable successful activation of target genes. This has also been demonstrated by others in recent studies. For example, Zhou and colleagues have reported that VPR strongly outperforms VP64 in activation of *Neurog2*, but *Ascl1* activation cannot be strongly

increased by strengthening effector (Zhou et al. 2018). Considering all of our experiments and work done by others, it is most likely that such failure is due to nucleosome positioning and presence of alternative TSS sites/promoter. Larger gRNA screen and use of chromatin remodelling enzymes should be tested to see if such hypothesis is correct.

6.1.5 Epigenetic modifications as a main limiting factor for efficient and robust sTF function.

From our and other group work some common limiting factors seem to emerge. Although we have not directly investigated nucleosome positioning in our experiments, there is an ample of studies which have directly showed that adding nucleosomes to DNA strands *in vitro* makes them inaccessible to Cas9 and dCas9 (Horlbeck et al. 2015; Isaac et al. 2014). Moreover, if nucleosome remodelling enzymes are used (such as γ Chd1, RSC or SNF2), Cas9 becomes capable of efficiently binding and cutting the target site. Interestingly, if distal parts (limer areas) of nucleosomes are targeted, Cas9 is able to cut DNA strand (Isaac et al. 2014). Considering that pioneering factors have an ability to remodel nucleosome positioning within hours of being expressed – as shown by Ascl1 ATAC-seq results hours after overexpression – it would be extremely beneficial to give such pioneering function to Cas9 by tethering it to appropriate remodelling enzyme (Wapinski et al. 2017). Although nucleosomes have been shown to be key for Cas9 binding, information regarding other epigenetic marks and their influence on Cas9 is unclear. For example, Hu and colleagues have reported that although TALE- and CRISPR-based transcriptional activation of Oct4 is possible in HEK293T cells, but DNA methylation that is found on the Oct4 promoter was not removed during this process

(J.Hu et al. 2014). However, more recent evidence shows that histone and DNA methylation can be altered by sTFs (Black et al. 2016). Further investigations will need to be undertaken to fully understand how DNA methylation might effect sTF functionality.

6.2 Toolkit for simultaneous gene activation and repression in multiplex

Our long-term goal remains to be ability to effectively activate and repress multiple target genes simultaneously in an individual cell. A number of sTF limitations have become obvious from the early experiments. Firstly, to activate genes in a reliable and robust manner a combination of gRNAs need to be used. It is difficult to predict this in advance; each must be screened to empirically determine the best activity in the right cell type. To test multiple gRNAs against multiple target genes simultaneously a new delivery system has to be developed; transient co-transfection would not suffice. Furthermore, gene repression would be highly desirable: not only does it play an important role directly in the cell fate reprogramming, but also repression of repressors might be essential for some master regulators to overcome biological constraints. For these reasons, we developed an all-in one (Ai1) system and VRER dCas9 mutant proteins.

6.2.1 An Extensible Mammalian Modular Assembly (EMMA), system for simple and rapid construction of expression plasmids

To robustly and homogeneously deliver multiple sTFs into cells, we needed to employ a cloning system which would allow rapid and modular assembly of many

parts with high efficiency. Expression plasmids need to be produced containing multiple unique gRNAs, plus a dCas9-Effector. With a help of Andrea Martella (Patrick Cai and S. Pollard Lab) we have constructed “EMMA: an Extensible Mammalian Modular Assembly” (Martella et al. 2017). EMMA is based on the Golden Gate cloning system, but is highly optimised for multipart assembly, with a large collection of validated mammalian parts. Using EMMA we have made expression vectors with various levels of complexity and tested them in HEK293T cells. This has demonstrated that complex expression vectors can be readily produced using this cloning strategy.

We went on to adapt this system for assembly of sTFs expression plasmids. Firstly, we constructed an empty Ai1 vector, which contains: cloning sites for gRNA insertion, CAG promoter-driven dCas9-VP160, PiggyBac homology arms for stable integration and a reporter mRuby gene. After validating functionality of its parts, we went on to test capacity for gRNA insertion.

6.2.2 Simultaneous activation of multiple endogenous genes by driving expression of three gRNAs-tRNA units from single U6 promoter

For the first set of sTFs, each gRNA was driven by a single U6 promoter. In addition to that, they also had gRNA scaffold fragments and termination sequences. This composition of sTF expression would generate a highly repetitive encoding DNA sequence if simply concatamerised, leading to recombination in the plasmid or cell. To avoid this anticipated problem, we decided to use a single U6 promoter which is followed by gRNAs separated by tRNA sequences (Xie et al. 2015). Transcription produces a long gRNA-tRNA transcript, but the activity of RNase H and RNase P

excise the flanking tRNA sequence leading to release of functional gRNAs, which then can form complexes with dCas9-VP160. This approach would not only make the sTFs plasmid sequence less repetitive (if different tRNAs would be used), but also minimise the length of it by a third (as tRNA sequence is 81bp long while U6 promoter is 264bp long).

This work was inspired by another group which managed to construct U6-tRNA-gRNA and used it successfully to cut multiple genes with Cas9 in rice (Xie et al. 2015). However, at the time we constructed our first U6-(tRNA-gRNA)₁....-(tRNA-gRNA)_n, there was no published work in mammalian systems nor was it ever used to activate rather than cut genes. By the time of writing this thesis, tRNA-gRNA systems have been used in mammalian cell for gene editing but not for gene activation (Dong et al. 2017; Xu et al. 2017).

Our first gRNA-tRNA construct contained the Glu-tRNA sequence and three gRNAs targeting *Sox10*, *Olig2* and *Nkx6-2*. After transfecting MEFs and checking for activation of those genes, we have found that all three genes were activated. This proof of principle experiment has demonstrated not only that the tRNA-gRNA sequence are properly transcribed but also cut in a way that leaves remaining gRNAs functional.

After this early success, we decided to construct a more elaborate gRNA-tRNA sequences which contained 8 different gRNAs. After a number of cloning efforts, we could not successfully make such a plasmid. We assumed this was because construct sequence was too repetitive (causing homologous recombination to occur during cloning process) and therefore switched to an original approach of using multiple U6

promoters. In the future, this problem could be solved by using different tRNA sequences instead of just using Glu-tRNA in all of the constructs.

6.2.3 All-in one plasmid, driven by multiple U6 promoters, for simultaneous activation of multiple endogenous genes

A second method which we developed was based on putting multiple U6-gRNA repeats into empty Ai1 vector. By the time of writing this thesis, other groups have published impressive methods for development of all-in one plasmids for sTF delivery. For example, Shao and colleagues have constructed all-in one plasmid containing 20 gRNAs targeting 20 different genes. They have managed to activate and repress 10 genes separately using such plasmid (Shao et al. 2018). To construct such multiplexing capacity, we started by adding three repeats which contained three different gRNAs targeting *Sox10*, *Olig2* and *Nkx6-2* (placed in that order). After successful activation of targeted genes in MEFs, we built larger plasmid which contained 2 gRNAs for *Sox10*, 2 gRNAs for *Olig2* and 1 gRNA for *Nkx6-2* (2SON). This again worked in MEFs and gave larger activation of *Sox10* and *Olig2*, but slightly lower average activation of *Nkx6-2*. A third Ai1 plasmid was constructed, termed 3SON, and it contained eight U6-gRNA repeats: 3 gRNAs targeting *Sox10*, 3 gRNAs targeting *Olig2* and 2 gRNAs targeting *Nkx6-2*. These genes could all be activated from this plasmid. However, *Olig2* and *Nkx6-2* had a lower level of activation compared to 2SON. This suggests that the longer Ai1 construct is, the lower activation from distal U6-gRNA units can be observed – although more variations and testing would be needed to confirm this. We speculate that there might be transcriptional competition or interference from surrounding U6 promoters. Future experiments might make use of addition of insulators or larger linker spaces between individual U6-gRNA

units. However, for the purpose of this thesis, not only have we shown that construction of such Ai1 plasmids is possible but also that they are able to activate multiple genes from the single expression vector delivering multiple guides for each gene.

6.2.4 VRER mutant dCas9 recognizes NGCG instead of NGG PAM

A second problem that we have set out to tackle was construction of a sTF system which would allow activation and repression of different genes using dCas9 simultaneously. A simple CRISPR/dCas9 system would be incapable of segregating repressive and activating functions as both dCas9 proteins would compete equally for gRNAs – a mixture of activators and repressors for each promoter/gene would emerge. To be able to direct dCas9-Activator and dCas9-Repressor to distinct target genes in the same cell dCas9 with distinct PAM sequence recognition would be needed. This could either be achieved by using different orthologues of dCas9 (from different bacterial species) or by protein engineering to mutate dCas9 protein to adjust its PAM preference (Kleinstiver et al. 2015). Another approach was recently reported that enables different effectors to be bound to the gRNA via use of RNA-Protein scaffolds (Konermann et al. 2014). This system is called SAM, and it allows construction of activating and repressive sTFs which can work in same cells. We have decided to choose first approach and mutate dCas9 protein, adjusting its PAM specificity in such a manner. Although this has been done with Cas9 (creating a VRER-Cas9 mutant) to show almost exclusive PAM specificity, it has not yet been used with dCas9 for transcriptional activation and repression, or multiplexed within the same cell.

We successfully introduced four amino acid substations (VRER) into dCas9 to produce the VRER-dCas9-VP160 and VRER-dCas9-KRAB versions of it. To test if VRER-dCas9-VP160 recognizes NGCG PAM instead of original NGG PAM, we have created gRNAs for both types of PAM, targeting *Nkx6-2* and *Sox10*. When targeting *Nkx6-2* in MEFs, we observed expected results: VRER-dCas9-VP160 could only activate *Nkx6-2* if ‘NGCG’ binding gRNAs were used and not if ‘NGG’ were used. To a large extent the opposite was true for dCas9-VP160. That being said, this was not quite so stringent, as dCas9-VP160 was still able to slightly activate *Nkx6-2* gene transcription (around 20-fold change) with the ‘wrong’ NGCG gRNAs. However, this was significantly lower compared to NGG gRNAs (around 4000-fold different). To some extent this agrees with results presented by authors of original study as they also saw small activity cross-PAMs. For example, recent work by Hu and colleagues have reported similar trends (J.Hu et al. 2018). They have used continues evolution in *E.coli* cells to force new mutations in Cas9 (and dCas9) protein and some of these were recognizing new PAM sequences. After direct head-to-head comparison of various mutants, they have found that some cross-reactivity remains (gRNAs with wrong PAM are able of activation). Although, as authors claim, such dCas9 characteristic is an important development for expanding target repertoire, but this might also have a negative effect on use of these reagents for simultaneous gene activation and repression as some competition for gRNAs will be present. However, it is likely that the specificity will be strong enough for cross-PAM recognition not to have a significant effect.

Our results were not so promising for *Sox10*. dCas9-VP160 managed to activate *Sox10* transcription if gRNAs containing the NGG PAM (original) were used,

but not if NGCG (new) PAMs were used. However, none of the gRNAs could activate *Sox10* gene with the mutant VRER-dCas9-VP160. While we would not expect NGG gRNAs to work, it was strange that NGCG gRNAs did not work as well. This may be because NGCG gRNAs could not be designed to completely target same promoter base pairs as NGG gRNAs. Considering that even small positional shift of gRNA location might make it completely non-functional – mainly because of nucleosome positioning but also because of possible secondary structure formation - we hypothesize that this is a reason for such results.

We went on to test if we can repress genes using VRER-dCas9-KRAB. For candidate gene targets, we have chosen *Sox9* and *Olig2*. In both case we managed to see a drop in positive cells after transfection with VRER-dCas9-KRAB and corresponding gRNAs. Although this was evident at protein level in a number of biological replicates, we could not observe any difference at mRNA level. We are unsure the explanation for this observation, but similar trend has been shown by a previous PhD student in our lab as well (S. Gogolok). One explanation could be related to the extent of the repression. Firstly, instead of observing more negative cells, we found more cells which are low in either of the proteins (as measured by the intensity of immunostaining). This means transcription is not completely turned off but rather slowed down. Secondly, this shift was observed only in small proportion of cells (less than 15%). Taken together, these results might explain why we see difference at single cell level but this signal is not significant enough when assessed across the whole cell population.

6.3 Programing and reprograming NSC and MEFs using synthetic transcription factors

In recent years, a number of groups have reported successful programing and reprograming of cell fate using sTFs. The best and most impressive example to date is work done by Black and colleagues. They have demonstrated that difficult direct lineage conversion from MEFs into neurons can be facilitated by using 12 gRNAs targeting *Ascl1*, *Brn2* and *Myt1l* gene transcription (Black et al. 2016). Although this was impressive study demonstrating strong potential of sTFs in cell fate programming, a number of features still have to be improved. In this study, multiple plasmids have been used which results in heterogeneous delivery of sTFs potentially compromising cell programming efficiency (considering each element is essential for sTF functionality). Furthermore, lentiviral-based delivery systems have been used which have limited cargo capacity. To target higher number of genes (which might assist with lineage conversion dynamics), plasmid or artificial chromosome-based delivery will need to be used.

For the last part of this work, we have tested if sTFs could have any biological effect on cell differentiation or phenotype – both NS cell programing to OPCs and MEF transdifferentiation towards OPCs/OLs. We have observed functional changes in both MEF and NSC after target gene activation using sTFs. In this work, we have used all-in one plasmids which are capable of delivering at least 8 different gRNAs and dCas9-VP160 within single plasmid.

6.3.1 *Sox10* overexpression or activation drives NSC differentiation towards OL lineage

To test if *Sox10* overexpression or activation has any effect on NSC differentiation, we have constructed stable cell lines harbouring appropriate cDNA overexpression or gene activation (sTF-based) system. Firstly, to have positive control in differentiation experiments we have established a cell line which was upregulating *SOX10* and *OLIG2* expression in presence of DOX. We had a number of problems minimizing leakage of these two genes in absence of DOX, and although we managed to find clones which were not significantly leaky, some basal expression of *SOX10* was still observed. Similarly, we constructed NSC lines which were constitutively activating *Sox10* due to expression of the dCas9-VP160 based sTFs.

Firstly, we tested if PDGFR α -GFP expression could be induced in NS cells under self-renewal conditions and if this could be enhanced to higher levels during differentiation. Under self-renewal conditions, cells which had *SOX10* upregulated (either by cDNA overexpression or sTF-based activation) were able to generate increased numbers of PDGFR α positive cells. Similar increases were observed when using differentiation inducing conditions. Thus, cells with activated *Sox10* gene (sTF) or overexpressed *SOX10* cDNA were able to generate significantly more PDGFR α -GFP cells at every time-point.

There are couple of important caveats for these results. Firstly, we have observed that all cell lines which had high *SOX10* were also proliferating faster compared to parental cells. To overcome this problem during our experiments, we have seeded cells at very low density (20,000 cells per 6-well plate at Day 0) which kept

cells at low confluence. In such a way, the effect on differentiation was coming more directly from gene regulation rather than changes in density. Secondly, PDGFR α is a marker of early stage OPCs rather than OLs. Moreover, SOX10 directly controls PDGFR α expression. Because of these reasons, described experiments on PG1.1 NSC cell line were not conclusive enough to state that *Sox10* is able to enhance NSC differentiation to OL lineage. At best, we could speculate that *Sox10* overexpression/activation was priming NSC to become OPCs.

To test if *Sox10* transcription activation (by sTFs) can enhance NSC differentiation to OLs, we have repeated the experiments on a different non-manipulated NSC line, called BL6, and forced them to differentiate fully to OLs. In the initial experiments, we have used a two-step differentiation protocol (FGF-2, PDGFR-AA and Forskolin followed by a switch to T3 and Ascorbic acid) and they resulted in a high frequency of O4 positive cells. Using this protocol, we could not see any differences between cells which had *Sox10* gene activated and cells which did not (both gave about 60% O4 positive cells). However, when using a sub-optimal differentiation protocol (addition of 5% FCS) we observed a dramatic increase in the levels of O4 positive cells comparing with and without the *Sox10* activating sTF. Activation of *Sox10* transcription by sTFs gave ~10% more O4 positive cells in these conditions. Further work would need to be undertaken to fully understand the genes required to force NSC differentiation under different conditions (i.e. there still seems to be some rate limiting event that limits total numbers of possible O4 positive cells specified). Thus, we demonstrated a significant impact on differentiation programs using sTFs in NSCs.

6.3.2 *Sox10*, *Olig2* and *Nkx6-2* activation by sTFs drive MEF transdifferentiation towards OL lineage

The second use of sTFs we explored for cell fate programming was to apply sTFs to drive MEF transdifferentiation to OPCs/OLs.

Two key published studies have shown that *Sox10*, *Olig2* and *Nkx6-2* or *Zfp653* overexpression in MEFs can trigger a subset of cells to become OL-like cells. These cells were positive for a number of OL markers; displayed a typical oligodendrocyte morphology; they were able to generate myelin sheets when transplanted into mice. In each of these studies the protocols used were similar, but not identical: cells were infected with the cDNAs expressed from viral vectors, and then maintained in MEF medium for 3 days. These were then seeded into Laminin coated plates containing transdifferentiation inducing medium. Media composition varied between two key studies. 14 to 24 days after TD-inducing medium addition, cells were scored for transdifferentiation events by flow cytometry and immunostaining for OL markers O4 and PLP1. Although the authors claimed that it is difficult to know the exact transdifferentiation efficiency (as many cells replicate and die throughout the process), but they have reported between 4-36% (17% on average) of O4 positive cells at day 21 in one study (Yang et al) and around 9% of O4 positive cells at day 24 in another study (Najm *et al*).

We have set out to repeat these findings by using our sTFs which are able to activate *Sox10*, *Olig2* and *Nkx6-2* in MEFs. For our experiments, we have used protocol described by Najm *et al*. Cells were transfected with sTFs at Day 0. All three genes were activated in MEFs at Day 3, 9 and 14 as shown by qPCR. However, at Day

17 we only saw 3.2% of O4 positive cells after using 3SON sTFs. After increasing the transdifferentiation protocol from 17 days to 24 days, we have found that this percentage increased to 6.9%.

There are couple important points requiring further discussion. Firstly, the level of gene activation seems to play a key role in transdifferentiation. In previous studies, couple thousand-fold change in gene transcription has been seen while our 3SON sTFs were able to activate *Sox10* to couple hundred-fold at most. Interestingly, supplementary data from Najm *et al* study shows that if any of three factors are missing, amount of PLP positive cells at day 21 drops to 4.7 – 7.6%. Following these observations, we decided to strength our sTFs by additional sTFs targeting *Sox10* gene (S10 vector). By doing so, we managed to increase the proportion of O4-positive cells from 6.9% to 17.9%. To even further enhance the effect, stronger effectors could be used as well. Secondly, length of transdifferentiation protocol plays important role. In the study by Yang *et al*, they have seen only 2.0 – 8.5% of cells expressed PLP at day 7. Furthermore, recent study which replicated findings from first Najm *et al* and Yang *et al* have reported even longer times required for transdifferentiation. In one case, they have left MEFs for over 30 days (Lee & Park 2017). We have also observed similar trends. Last thing which is important to mention is transdifferentiation medium composition. In our studies, we have used medium described by Najm *et al*. It would be interesting to see if using medium defined by Yang *et al* would make any difference.

6.4 Future work

Here we report an effort to build and screen multiple sTFs for the purpose of activating *Sox10*, *Olig2*, *Nkx6-2* and *Nkx2-2* in neural stem cells and mouse embryonic

fibroblasts. We have tested these sTFs in different contexts including NS cell self-renewal and differentiation as well as MEF transdifferentiation induction. Furthermore, we have designed and built tools for multiple gene modulation from a single plasmid as well as mutant sTF variants which allows to adjust target specificity. Lastly, we investigated how multiple sTF cooperate to activate target genes.

To summarize, we have successfully achieved the following. We have activated *Sox10* gene in NS using dCas9-VP160 and multiple gRNAs. After investigating potency of these cells to differentiate into OPCs and OLs, we showed that *Sox10* transcription activation allows more cells to differentiate to OL-lineage. Next, we have built a system which allows to generate expression plasmids containing dCas9-VP160 and up to eight different gRNAs in the same plasmid. This allowed activation of three different genes (*Sox10*, *Olig2* and *Nkx6-2*) from a single plasmid in MEFs. Furthermore, we have built and validated dCas9-VP160 mutant, VRER, which has modified PAM specificity and therefore searches for alternative target sites. Such mutant was almost exclusively recognizing gRNAs with NGCG PAM while ignoring its natural targets with NGG PAM. Next, we show preliminary single-cell results demonstrating that using multiple gRNAs to activate *OLIG2* in HEK293T cells not only allows more cells to be activated but also activation level within these cells is higher as compared to using single gRNA. Lastly, we have shown that sTF-driven activation of *Sox10*, *Olig2* and *Nkx6-2* endogenous transcription can drive MEF transdifferentiation to O4-positive OPC-like cells in 24 days.

While our work helped to better understand sTFs and their potential role in cell fate reprogramming, there are still many things which need to be further validated and/or optimized. Some of the directions for future work we will undertake includes (i) better

understanding what impact different TADs have on gene activation at single cell level, (ii) optimizing the MEF transdifferentiation protocol and confirming that functional oligodendrocytes can be produced. Other important areas which needs further investigation include: testing multiplexing limits for Ai1 expression vectors; investigating potential of using different tRNAs in Ai1 expression vectors; investigating the SAM system for gene repression; testing multiplexing system for simultaneous gene activation and repression. Altogether the sTF technologies show great promise for controlling endogenous gene activation and repression. In the future these will likely provide important tools and reagents for steering mammalian cell fate in vitro and in vivo.

7 References

- Acar, M., Becskei, A. & van Oudenaarden, A., 2005. Enhancement of cellular memory by reducing stochastic transitions. *Nature*, 435(7039), pp.228–232. Available at: <http://www.nature.com/doifinder/10.1038/nature03524> [Accessed August 16, 2017].
- Ahmed, S. et al., 2009. Transcription factors and neural stem cell self-renewal, growth and differentiation. *Cell adhesion & migration*, 3(4), pp.412–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19535895> [Accessed August 16, 2017].
- Arvidsson, A. et al., 2002. Neuronal replacement from endogenous precursors in the adult brain after stroke. *Nature Medicine*, 8(9), pp.963–970. Available at: <http://www.nature.com/doifinder/10.1038/nm747> [Accessed August 16, 2017].
- Balboa, D. et al., 2015. Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. *Stem Cell Reports*, 5(3), pp.448–459. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26352799> [Accessed August 16, 2017].
- Ben-Hur, T., 2011. Cell Therapy for Multiple Sclerosis. *Neurotherapeutics*, 8(4), pp.625–642. Available at: <http://link.springer.com/10.1007/s13311-011-0073-x> [Accessed August 16, 2017].
- Bikard, D. et al., 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Research*, 41(15), pp.7429–7437. Available at: <https://academic.oup.com/nar/article->

lookup/doi/10.1093/nar/gkt520 [Accessed August 16, 2017].

Black, J.B. et al., 2015. Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells. *Cell Stem Cell*, pp.1–9. Available at: <http://dx.doi.org/10.1016/j.stem.2016.07.001>.

Boch, J. et al., 2009. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, 326(5959), pp.1509–1512. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19933107> [Accessed August 16, 2017].

Bolotin, A. et al., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 151(8), pp.2551–2561. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16079334> [Accessed August 16, 2017].

Bonaguidi, M.A. et al., 2005. LIF and BMP signaling generate separate and discrete types of GFAP-expressing cells. *Development*, 132(24). Available at: <http://dev.biologists.org/content/132/24/5503.long> [Accessed August 16, 2017].

Braun, S.M.G. et al., 2015. Programming Hippocampal Neural Stem/Progenitor Cells into Oligodendrocytes Enhances Remyelination in the Adult Brain after Injury. *Cell reports*, 11(11), pp.1679–85.

Budde, H. et al., 2010. Control of oligodendroglial cell number by the miR-17-92 cluster. *Development*, 137(13), pp.2127–2132. Available at: <http://dev.biologists.org/cgi/doi/10.1242/dev.050633> [Accessed August 16, 2017].

- Buganim, Y., Faddah, D.A. & Jaenisch, R., 2013. Mechanisms and models of somatic cell reprogramming. *Nature reviews. Genetics*, 14(6), pp.427–39. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23681063> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4060150> [Accessed August 16, 2017].
- Cai, J. et al., 2010. Co-localization of Nkx6.2 and Nkx2.2 homeodomain proteins in differentiated myelinating oligodendrocytes. *Glia*, 58(4), pp.458–68.
- Cai, J. et al., 2005. Generation of oligodendrocyte precursor cells from mouse dorsal spinal cord independent of Nkx6 regulation and Shh signaling. *Neuron*, 45(1), pp.41–53.
- Carninci, P. et al., 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics*, 38(6), pp.626–635. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16645617> [Accessed August 16, 2017].
- Cermak, T. et al., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research*, 39(12), pp.e82–e82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21493687> [Accessed August 16, 2017].
- Chakraborty, S. et al., 2014. A CRISPR/Cas9-based system for reprogramming cell lineage specification. *Stem Cell Reports*, 3(6), pp.940–947. Available at: <http://dx.doi.org/10.1016/j.stemcr.2014.09.013> [Accessed August 16, 2017].
- Chavez, A. et al., 2016. Comparison of Cas9 activators in multiple species. *Nat Methods*, 13(7), pp.563–567.

- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer, E.P.R., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., Dmitry, T.-O., Braff, J.L., Davidsohn, N., Housden, B.E., Perrimon, N., Weiss, R., Aach, J., Collins, J.J., Church, G.M., et al., 2015. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*, 12(4), pp.326–328. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25730490> [Accessed September 7, 2016].
- Chen, B. et al., 2013. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, 155(7), pp.1479–1491. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24360272> [Accessed August 16, 2017].
- Cho, S.W. et al., 2014. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research*, 24(1), pp.132–41. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24253446> [Accessed August 16, 2017].
- Cong, L. Le et al., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339(6121), pp.819–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23287718> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3795411> [Accessed September 6, 2016].
- Conti, L. et al., 2005. Niche-Independent Symmetrical Self-Renewal of a Mammalian Tissue Stem Cell R. Lovell-Badge, ed. *PLoS Biology*, 3(9), p.e283. Available at: <http://dx.plos.org/10.1371/journal.pbio.0030283> [Accessed October 7, 2017].

- Cooper, S.J. et al., 2005. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Research*, 16(1), pp.1–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16344566> [Accessed August 16, 2017].
- Copray, S. et al., 2006. Olig2 Overexpression Induces the In Vitro Differentiation of Neural Stem Cells into Mature Oligodendrocytes. *STEM CELLS*, 24(4), pp.1001–1010. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16253982> [Accessed August 16, 2017].
- Davis, R.L., Weintraub, H. & Lassar, A.B., 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*, 51(6), pp.987–1000. Available at: <http://www.sciencedirect.com/science/article/pii/009286748790585X?via%3Di> hub [Accessed October 3, 2017].
- Dong, F. et al., 2017. Polycistronic tRNA and CRISPR guide-RNA enables highly efficient multiplexed genome engineering in human cells. *Biochemical and Biophysical Research Communications*, 482(4), pp.889–895. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27890617> [Accessed November 15, 2017].
- Douvaras, P. et al., 2014. Efficient generation of myelinating oligodendrocytes from primary progressive multiple sclerosis patients by induced pluripotent stem cells. *Stem Cell Reports*, 3(2), pp.250–259. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25254339> [Accessed August 16, 2017].

- Ebina, W. & Rossi, D.J., 2015. Transcription factor-mediated reprogramming toward hematopoietic stem cells. *The EMBO Journal*, 34(6), pp.694–709. Available at: <http://emboj.embopress.org/cgi/doi/10.15252/emboj.201490804> [Accessed August 16, 2017].
- Emery, B. et al., 2009. Myelin Gene Regulatory Factor Is a Critical Transcriptional Regulator Required for CNS Myelination. *Cell*, 138(1), pp.172–185. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19596243> [Accessed August 16, 2017].
- Emery, B., 2010. Regulation of Oligodendrocyte Differentiation and Myelination. *Science*, 330(6005), pp.779–782. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21051629> [Accessed August 16, 2017].
- Eriksson, P.S. et al., 1998. Neurogenesis in the adult human hippocampus. *Nature Medicine*, 4(11), pp.1313–1317. Available at: <http://www.nature.com/doi/10.1038/3305> [Accessed August 16, 2017].
- Feng, B. et al., 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature Cell Biology*, 11(2), pp.197–203. Available at: <http://www.nature.com/doi/10.1038/ncb1827> [Accessed August 16, 2017].
- Fidanza, A. et al., 2017. An all-in-one UniSam vector system for efficient gene activation. *Scientific reports*, 7(1), p.6394. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28743878> [Accessed October 8, 2017].
- Fishman, V.S.V. et al., 2015. Cell divisions are not essential for the direct conversion

of fibroblasts into neuronal cells. *Cell Cycle*, 14(8), pp.1188–1196. Available at:
<http://www.tandfonline.com/doi/full/10.1080/15384101.2015.1012875>
[Accessed August 16, 2017].

Franco, P.G., Pasquini, J.M. & Silvestroff, L., 2015. Optimizing culture medium composition to improve oligodendrocyte progenitor cell yields in vitro from subventricular zone-derived neural progenitor cell neurospheres. *PloS one*, 10(4), p.e0121774. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/25837625> [Accessed August 16, 2017].

Fujita, T. & Fujii, H., 2013. *Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR*, Available at: <http://www.sciencedirect.com/science/article/pii/S0006291X13013296>
[Accessed August 16, 2017].

Gasiunas, G. et al., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, 109(39), pp.E2579–E2586. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/22949671> [Accessed August 16, 2017].

Gersbach, C. a & Perez-Pinera, P., 2014. Activating human genes with zinc finger proteins, transcription activator-like effectors and CRISPR/Cas9 for gene therapy and regenerative medicine. *Expert opinion on therapeutic targets*, 18(November 2015), pp.1–5. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/24917359> [Accessed August 16, 2017].

- Ghazy, M.A. et al., 2004. Amino Acid Substitutions in Yeast TFIIF Confer Upstream Shifts in Transcription Initiation and Altered Interaction with RNA Polymerase II. *Molecular and Cellular Biology*, 24(24), pp.10975–10985. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15572698> [Accessed August 16, 2017].
- Gilbert, L.A.A. et al., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159(3), pp.647–661. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25307932> [Accessed August 16, 2017].
- Givogri, M.I. et al., 2006. Oligodendroglial Progenitor Cell Therapy Limits Central Neurological Deficits in Mice with Metachromatic Leukodystrophy. *Journal of Neuroscience*, 26(12), pp.3109–3119. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16554462> [Accessed August 16, 2017].
- Glaser, T. et al., 2007. Tripotential Differentiation of Adherently Expandable Neural Stem (NS) Cells M. van Ijzendoorn, ed. *PLoS ONE*, 2(3), p.e298. Available at: <http://dx.plos.org/10.1371/journal.pone.0000298> [Accessed September 6, 2016].
- Golding, I. et al., 2005. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6), pp.1025–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16360033> [Accessed August 16, 2017].
- Groner, A.C. et al., 2010. KRAB–Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading H. D. Madhani, ed. *PLoS Genetics*, 6(3), p.e1000869. Available at: <http://dx.plos.org/10.1371/journal.pgen.1000869> [Accessed September 7, 2016].
- Hall, D.B. & Struhl, K., 2002. The VP16 Activation Domain Interacts with Multiple

Transcriptional Components as Determined by Protein-Protein Cross-linking *in Vivo*. *Journal of Biological Chemistry*, 277(48), pp.46043–46050. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12297514> [Accessed August 16, 2017].

Hamilton, T.G. et al., 2003. Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. *Molecular and cellular biology*, 23(11), pp.4013–25. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12748302> [Accessed October 9, 2017].

Hanna, J. et al., 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*, 462(7273), pp.595–601. Available at: <http://www.nature.com/doifinder/10.1038/nature08592> [Accessed August 16, 2017].

Hardwick, J.M. et al., 1992. The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *Journal of virology*, 66(9), pp.5500–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1323708> [Accessed September 7, 2016].

Hilton, I.B. et al., 2015. Epigenome editing by a {CRISPR-Cas9-based} acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*, 33(5), pp.510–517.

Himeda, C.L., Jones, T.I. & Jones, P.L., 2016. CRISPR/dCas9-mediated Transcriptional Inhibition Ameliorates the Epigenetic Dysregulation at D4Z4 and Represses DUX4-fl in FSH Muscular Dystrophy. *Molecular Therapy*, 24(3), pp.527–535. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/26527377> [Accessed September 7, 2016].

Hinz, J.M., Laughery, M.F. & Wyrick, J.J., 2015. Nucleosomes Inhibit Cas9 Endonuclease Activity in Vitro. *Biochemistry-us*, 54(48), pp.7063–7066.

Hirai, H., Tani, T. & Kikyo, N., 2010. Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *The International Journal of Developmental Biology*, 54(11–12), pp.1589–1596. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21404180> [Accessed August 16, 2017].

Horlbeck, M.A. et al., 2016. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife*, 5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26987018> [Accessed August 16, 2017].

Hou, P. et al., 2013. Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. *Science*, 341(6146). Available at: <http://science.sciencemag.org/content/341/6146/651.long> [Accessed August 16, 2017].

Hu, J. et al., 2014. Direct activation of human and mouse Oct4 genes using engineered {TALE} and Cas9 transcription factors. *Nucleic Acids Res*, 42(7), pp.4375–4390.

Hu, J. et al., 2014. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, 556(7699), pp.57-63.

Huisinga, K.L. & Pugh, B.F., 2004. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Molecular cell*, 13(4), pp.573–85. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/14992726> [Accessed August 16, 2017].

Ikawa, T. et al., 2015. Induced Developmental Arrest of Early Hematopoietic

Progenitors Leads to the Generation of Leukocyte Stem Cells. *Stem Cell*

Reports, 5(5), pp.716–727. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/26607950> [Accessed August 16, 2017].

Imbalzano, A.N., Zaret, K.S. & Kingston, R.E., 1994. Transcription factor (TF) IIB

and TFIIA can independently increase the affinity of the TATA-binding protein

for DNA. *The Journal of biological chemistry*, 269(11), pp.8280–6. Available

at: <http://www.ncbi.nlm.nih.gov/pubmed/8132551> [Accessed August 16, 2017].

Isaac, R.S. et al., 2016. Nucleosome breathing and remodeling constrain CRISPR-

Cas9 function. *eLife*, 5. Available at:

<http://elifesciences.org/lookup/doi/10.7554/eLife.13450> [Accessed August 16,

2017].

Iwafuchi-Doi, M. et al., 2014. Pioneer transcription factors in cell reprogramming.

Gene Dev, 28(24), pp.2679–2692. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/25512556> [Accessed September 7, 2016].

Iwafuchi-Doi, M. & Zaret, K.S., 2016. Cell fate control by pioneer transcription

factors. *Development*, 143(11). Available at:

<http://dev.biologists.org/content/143/11/1833.long> [Accessed August 16, 2017].

J.B., G., 1962. Adult frogs derived from the nuclei of single somatic cells.

Developmental Biology, 4(2), pp.256–273. Available at:

<http://www.sciencedirect.com/science/article/pii/001216066290043X> [Accessed

October 3, 2017].

Jarriault, S., Schwab, Y. & Greenwald, I., 2008. A *Caenorhabditis elegans* model for epithelial-neuronal transdifferentiation. *Proceedings of the National Academy of Sciences*, 105(10), pp.3790–3795. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/18308937> [Accessed October 3, 2017].

Jinek, M. et al., 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096), pp.816–821. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/22745249> [Accessed November 14, 2017].

Khalil, A.S. & Collins, J.J., 2010. Synthetic biology: applications come of age. *Nat Rev Genet*, 11(5), pp.367–379. Available at:

<http://www.nature.com/doifinder/10.1038/nrg2775> [Accessed August 16, 2017].

Kikuchi, K., 2015. Dedifferentiation, Transdifferentiation, and Proliferation:

Mechanisms Underlying Cardiac Muscle Regeneration in Zebrafish. *Curr Pathobiology Reports*, 3(1), pp.81–88. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/25722956> [Accessed August 16, 2017].

Kim, D. et al., 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nature Methods*, 12(3), pp.237–243. Available at:

<http://www.nature.com/doifinder/10.1038/nmeth.3284> [Accessed August 16, 2017].

Kim, H.J. et al., 2009. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Research*, 19(7),

pp.1279–1288. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19470664>
[Accessed August 16, 2017].

Kim, J.-S., 2016. Genome editing comes of age. *Nature Protocols*, 11(9), pp.1573–1578. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27490630> [Accessed August 16, 2017].

Kim, Y.G., Cha, J. & Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93(3), pp.1156–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8577732> [Accessed September 5, 2016].

Kleinstiver, B.P. et al., 2015. Engineered {CRISPR-Cas9} nucleases with altered {PAM} specificities. *Nature*, 523(7561), pp.481–485.

Konermann, S. et al., 2014. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536), pp.583–588. Available at: <http://www.nature.com/doifinder/10.1038/nature14136> [Accessed November 14, 2017].

Kopp, J.L. et al., 2012. Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*, 22(6), pp.737–750. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1535610812004497> [Accessed October 3, 2017].

Kuypers, N.J. et al., 2016. Remyelinating Oligodendrocyte Precursor Cell miRNAs from the Sfmbt2 Cluster Promote Cell Cycle Arrest and Differentiation. *The*

- Journal of neuroscience : the official journal of the Society for Neuroscience*, 36(5), pp.1698–710. Available at:
<http://www.jneurosci.org/content/36/5/1698.long> [Accessed August 16, 2017].
- Lander, E.S., 2016. The Heroes of CRISPR. Available at:
<http://dx.doi.org/10.1016/j.cell.2015.12.041> [Accessed November 14, 2017].
- Lee, E.-H. & Park, C.-H., 2017. Comparison of Reprogramming Methods for Generation of Induced-Oligodendrocyte Precursor Cells. *Biomolecules & therapeutics*, 25(4), pp.362–366. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/28605832> [Accessed November 16, 2017].
- Li, Q. et al., 2015. Transplantation of induced pluripotent stem cell-derived renal stem cells improved acute kidney injury. *Cell & bioscience*, 5, p.45. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26294957> [Accessed August 16, 2017].
- Li, Y. & Weiss, R., 2017. A Modular Approach to Building Complex Synthetic Circuits. In Humana Press, New York, NY, pp. 231–248. Available at:
http://link.springer.com/10.1007/978-1-4939-7223-4_17 [Accessed August 16, 2017].
- Lin, S. et al., 2015. In vivo transcriptional activation using CRISPR/Cas9 in *Drosophila*. *Genetics*, 201(2), pp.433–442.
- Lin, Y. et al., 2015. Combinatorial gene regulation by modulation of relative pulse timing. *Nature*, 527(7576), pp.54–58. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/26466562> [Accessed August 16, 2017].

Liu, Z. et al., 2007. Induction of oligodendrocyte differentiation by Olig2 and Sox10:

Evidence for reciprocal interactions and dosage-dependent mechanisms.

Developmental Biology, 302(2), pp.683–693. Available at:

<http://linkinghub.elsevier.com/retrieve/pii/S0012160606012978> [Accessed

August 16, 2017].

Llorens-Bobadilla, E. et al., 2015. Single-Cell Transcriptomics Reveals a Population

of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell*

Stem Cell, 17(3), pp.329–340. Available at:

<http://www.sciencedirect.com/science/article/pii/S193459091500301X>

[Accessed August 16, 2017].

Lois, C. & Alvarez-Buylla, A., 1993. Proliferating subventricular zone cells in the

adult mammalian forebrain can differentiate into neurons and glia. *Proceedings*

of the National Academy of Sciences of the United States of America, 90(5),

pp.2074–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8446631>

[Accessed August 16, 2017].

Lu, Q.R. et al., 2002. Common developmental requirement for Olig function

indicates a motor neuron/oligodendrocyte connection. *Cell*, 109(1), pp.75–86.

Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11955448> [Accessed

August 16, 2017].

Lupo, A. et al., 2013. KRAB-Zinc Finger Proteins: A Repressor Family Displaying

Multiple Biological Functions. *Current genomics*, 14(4), pp.268–78. Available

at: <http://www.ncbi.nlm.nih.gov/pubmed/24294107> [Accessed September 7, 2016].

Maire, C.L., Buchet, D., Kerninon, C., Deboux, C., Baron-Van Evercooren, A., Nait-Oumesmar, B., et al., 2009. Directing human neural stem/precursor cells into oligodendrocytes by overexpression of Olig2 transcription factor. *J Neurosci Res*, 87(15), pp.3438–3446. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19739249> [Accessed August 16, 2017].

Maire, C.L., Buchet, D., Kerninon, C., Deboux, C., Baron-Van Evercooren, A. & Nait-Oumesmar, B., 2009. Directing human neural stem/precursor cells into oligodendrocytes by overexpression of Olig2 transcription factor. *Journal of Neuroscience Research*, 87(15), pp.3438–3446. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19739249> [Accessed August 16, 2017].

Mali, P., Yang, L., et al., 2013. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339(6121), pp.823–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23287722> [Accessed September 7, 2016].

Mali, P., Esvelt, K.M. & Church, G.M., 2013. Cas9 as a versatile tool for engineering biology. *Nature Methods*, 10(10), pp.957–963. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24076990> [Accessed October 7, 2017].

Martella, A. et al., 2017. EMMA: An Extensible Mammalian Modular Assembly Toolkit for the Rapid Design and Production of Diverse Expression Vectors. *ACS Synthetic Biology*, 6(7), pp.1380–1392. Available at: <http://pubs.acs.org/doi/abs/10.1021/acssynbio.7b00016> [Accessed November

14, 2017].

Massirer, K.B. et al., 2011. Maintenance and differentiation of neural stem cells.

Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 3(1), pp.107–114. Available at: <http://doi.wiley.com/10.1002/wsbm.100> [Accessed August 16, 2017].

Mertens, J. et al., 2016. Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience. *Nat Rev Neurosci*, 17(7), pp.424–437. Available at: <http://dx.doi.org/10.1038/nrn.2016.46> [Accessed August 16, 2017].

Mojica, F.J.M. et al., 2005. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, 60(2), pp.174–182. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15791728> [Accessed August 16, 2017].

Morrison, S.J., Shah, N.M. & Anderson, D.J., 1997. Regulatory Mechanisms in Stem Cell Biology. *Cell*, 88(3), pp.287–298. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S009286740081867X> [Accessed August 16, 2017].

Murry, C.E. & Keller, G., 2008. Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons from Embryonic Development. *Cell*, 132(4), pp.661–680. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S009286740800216X> [Accessed August 16, 2017].

- Najm, F.J. et al., 2011a. Rapid and robust generation of functional oligodendrocyte progenitor cells from epiblast stem cells. *Nature Methods*, 8(11), pp.957–962. Available at: <http://www.nature.com/doifinder/10.1038/nmeth.1712> [Accessed August 16, 2017].
- Najm, F.J. et al., 2013. Transcription factor-mediated reprogramming of fibroblasts to expandable, myelinogenic oligodendrocyte progenitor cells. *Nature Biotechnology*, 31(5), pp.426–433. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3678540&tool=pmc&rendertype=abstract> [Accessed June 21, 2017].
- Nakagawa, M. et al., 2007. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnology*, 26(1), pp.101–106. Available at: <http://www.nature.com/doifinder/10.1038/nbt1374> [Accessed August 16, 2017].
- Neman, J. & de Vellis, J., 2012. A method for deriving homogenous population of oligodendrocytes from mouse embryonic stem cells. *Developmental Neurobiology*, 72(6), pp.777–788. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22234971> [Accessed August 16, 2017].
- Neri, M. et al., 2010. Robust Generation of Oligodendrocyte Progenitors from Human Neural Stem Cells and Engraftment in Experimental Demyelination Models in Mice. *Plos One*, 5(4), p.e10145. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20405042> [Accessed September 7, 2016].
- Nielsen, S. et al., 2013. Mechanism of eukaryotic RNA polymerase III transcription

- termination. *Science (New York, N.Y.)*, 340(6140), pp.1577–80. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/23812715> [Accessed September 7, 2016].
- Nishida, H. et al., 2006. Histone H3 acetylated at lysine 9 in promoter is associated with low nucleosome density in the vicinity of transcription start site in human cell. *Chromosome Research*, 14(2), pp.203–211. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/16544193> [Accessed August 16, 2017].
- O'Malley, J. et al., 2013. High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature*, 499(7456), pp.88–91. Available at:
<http://www.nature.com/doifinder/10.1038/nature12243> [Accessed August 16, 2017].
- Papapetrou, E.P. et al., 2009. Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), pp.12759–64. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/19549847> [Accessed August 16, 2017].
- Pawlowski, M. et al., 2017. Inducible and Deterministic Forward Programming of Human Pluripotent Stem Cells into Neurons, Skeletal Myocytes, and Oligodendrocytes. *Stem Cell Reports*, 8(4), pp.803–812.
- Perez-Pinera, P. et al., 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nature Methods*, 10(10), pp.973–976. Available at:
<http://www.nature.com/doifinder/10.1038/nmeth.2600> [Accessed August 16, 2017].

- Polentes, J. et al., 2012. Human Induced Pluripotent Stem Cells Improve Stroke Outcome and Reduce Secondary Degeneration in the Recipient Brain. *Cell Transplantation*, 21(12), pp.2587–2602. Available at: <http://journals.sagepub.com/doi/10.3727/096368912X653228> [Accessed August 16, 2017].
- Polstein, L.R. et al., 2015. Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Research*, 25(8), pp.1158–1169. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26025803> [Accessed September 7, 2016].
- Pozniak, C.D. et al., 2010. Sox10 directs neural stem cells toward the oligodendrocyte lineage by decreasing Suppressor of Fused expression. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), pp.21795–800.
- Radzisheuskaya, A. et al., 2016. Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. *Nucleic Acids Research*, 44(18), p.gkw583. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27353328> [Accessed September 7, 2016].
- Raj, A. et al., 2006. Stochastic mRNA Synthesis in Mammalian Cells U. Schibler, ed. *PLoS Biology*, 4(10), p.e309. Available at: <http://dx.plos.org/10.1371/journal.pbio.0040309> [Accessed August 16, 2017].
- Raj, A. & Tyagi, S., 2010. Detection of Individual Endogenous RNA Transcripts In

Situ Using Multiple Singly Labeled Probes. In *Methods in enzymology*. pp. 365–386. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20580972> [Accessed October 8, 2017].

Raposo, A. et al., 2015. Ascl1 Coordinately Regulates Gene Expression and the Chromatin Landscape during Neurogenesis. *Cell Reports*, 10(9), pp.1544–1556.

Rebar, E.J. et al., 2002. Induction of angiogenesis in a mouse model using engineered transcription factors. *Nature Medicine*, 8(12), pp.1427–1432. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12415262> [Accessed August 16, 2017].

Reynolds, B. & Weiss, S., 1992. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science*, 255(5052). Available at: <http://science.sciencemag.org/content/255/5052/1707.long> [Accessed August 16, 2017].

Reyon, D. et al., 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), pp.460–465. Available at: <http://www.nature.com/doifinder/10.1038/nbt.2170> [Accessed August 16, 2017].

Richardson, W.D. et al., 2006. Oligodendrocyte wars. *Nature reviews. Neuroscience*, 7(1), pp.11–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16371946> [Accessed September 6, 2016].

Robinton, D.A. & Daley, G.Q., 2012. The promise of induced pluripotent stem cells in research and therapy. *Nature*, 481(7381), pp.295–305. Available at:

<http://www.nature.com/doifinder/10.1038/nature10761> [Accessed August 16, 2017].

Sainsbury, S., Bernecky, C. & Cramer, P., 2015. Structural basis of transcription initiation by RNA polymerase II. *Nature Reviews Molecular Cell Biology*, 16(3), pp.129–143. Available at: <http://www.nature.com/doifinder/10.1038/nrm3952> [Accessed August 16, 2017].

Sandelin, A. et al., 2007. Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nature Reviews Genetics*, 8(6), pp.424–436. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17486122> [Accessed August 15, 2017].

Sander, J.D. & Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, 32(4), pp.347–355. Available at: <http://www.nature.com/doifinder/10.1038/nbt.2842> [Accessed November 14, 2017].

Scott J. et al., 2014. Targeted genome regulation and modification using transcription activator-like effectors. *FEBS Journal*, 281(20), pp.4583–4597.

Schultz, D.C. et al., 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & Development*, 16(8), pp.919–932. Available at: <http://www.genesdev.org/cgi/doi/10.1101/gad.973302> [Accessed September 7,

2016].

Shao, S. et al., 2018. Multiplexed sgRNA Expression Allows Versatile Single Nonrepetitive DNA Labeling and Endogenous Gene Regulation. *ACS Synthetic Biology*, pp.176-186.

Shen, B. et al., 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*, 11(4), pp.399–402. Available at: <http://www.nature.com/doifinder/10.1038/nmeth.2857> [Accessed November 14, 2017].

Shi, Y. et al., 2008. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell stem cell*, 3(5), pp.568–74. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18983970> [Accessed August 16, 2017].

Shi, Z. et al., 2016. Conversion of Fibroblasts to Parvalbumin Neurons by One Transcription Factor, Ascl1, and the Chemical Compound Forskolin. *The Journal of biological chemistry*, 291(26), pp.13560–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27137935> [Accessed August 16, 2017].

Singh, A.M. et al., 2015. Cell-Cycle Control of Bivalent Epigenetic Domains Regulates the Exit from Pluripotency. *Stem cell reports*, 5(3), pp.323–36. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2213671115002118> [Accessed August 16, 2017].

Slaymaker, I.M. et al., 2016. Rationally engineered Cas9 nucleases with improved specificity. *Science*, 351(6268), pp.84–88. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/26628643> [Accessed August 16, 2017].

Smale, S.T. et al., 1998. The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. *Cold Spring Harbor symposia on quantitative biology*, 63, pp.21–31. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10384267> [Accessed August 16, 2017].

Sokolik, C. et al., 2015. Transcription Factor Competition Allows Embryonic Stem Cells to Distinguish Authentic Signals from Noise. *Cell Systems*, 1(2), pp.117–129. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2405471215000538> [Accessed September 6, 2016].

Soufi, A. et al., 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, 151(5), pp.994–1004. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23159369> [Accessed August 16, 2017].

Soufi, A. et al., 2015. Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. *Cell*, 161(3), pp.555–568. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25892221> [Accessed August 16, 2017].

Sparman, M. et al., 2009. Epigenetic reprogramming by somatic cell nuclear transfer in primates. *Stem cells (Dayton, Ohio)*, 27(6), pp.1255–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19489081> [Accessed August 16, 2017].

Sternberg, S.H. et al., 2014. DNA interrogation by the CRISPR RNA-guided

endonuclease Cas9. *Nature*, 507(7490), pp.62–67. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/24476820><http://www.nature.com/nature/journal/v507/n7490/pdf/nature13011.pdf> [Accessed August 16, 2017].

Sundberg, M. et al., 2010. Production and isolation of NG2+ oligodendrocyte precursors from human embryonic stem cells in defined serum-free medium. *Stem Cell Research*, 5(2), pp.91–103. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/20538536> [Accessed August 16, 2017].

Suzuki, Y. et al., 2001. Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. *EMBO reports*, 2(5), pp.388–93. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/11375929> [Accessed August 16, 2017].

Takahashi, K. et al., 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4), pp.663–676. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16904174> [Accessed August 16, 2017].

Tarlow, B.D. et al., 2014. Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell stem cell*, 15(5), pp.605–18. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25312494> [Accessed October 3, 2017].

Triezenberg, S.J. et al., 1996. Transcriptional Activation Domain of the Herpesvirus Protein VP16 Becomes Conformationally Constrained upon Interaction with Basal Transcription Factors. *Journal of Biological Chemistry*, 271(9), pp.4827–4837. Available at: <http://www.jbc.org/cgi/doi/10.1074/jbc.271.9.4827>

[Accessed September 7, 2016].

- Tsai, S.Q. & Joung, K.J., 2016. Defining and improving the genome-wide specificities of {CRISPR-Cas9} nucleases. *Nat Rev Genet*, 17(5), pp.300–312.
- Tsonis, P.A. et al., 2004. A newt's eye view of lens regeneration. *The International journal of developmental biology*, 48(8–9), pp.975–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15558488> [Accessed August 16, 2017].
- van Oevelen, C. et al., 2015. C/EBP α Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis. *Stem Cell Reports*, 5(2), pp.232–247. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2213671115001885> [Accessed August 16, 2017].
- Vierbuchen, T. et al., 2010. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, 463(7284), pp.1035–1041. Available at: <http://www.nature.com/doifinder/10.1038/nature08797> [Accessed August 16, 2017].
- Vierbuchen, T. & Wernig, M., 2011. Direct lineage conversions: unnatural but useful? *Nature biotechnology*, 29(10), pp.892–907. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21997635> [Accessed September 5, 2016].
- Wang, C. et al., 2015. A Simple {CRISPR/Cas9} System for Multiplex Genome Editing in Rice. *J Genet Genomics*, 42(12), pp.703–706.
- Wang, C. et al., 2015. High purity of human oligodendrocyte progenitor cells obtained from neural stem cells: Suitable for clinical application. *J Neurosci*

Meth, 240, pp.61–66. Available at:

<http://linkinghub.elsevier.com/retrieve/pii/S0165027014003859> [Accessed September 7, 2016].

Wang, H., La Russa, M. & Qi, L.S., 2016. CRISPR/Cas9 in Genome Editing and Beyond. *Annual Review of Biochemistry*, 85(1), pp.227–264. Available at: <http://www.annualreviews.org/doi/10.1146/annurev-biochem-060815-014607> [Accessed August 16, 2017].

Wang, J., Pol, S.U., et al., 2014. Transcription factor induction of human oligodendrocyte progenitor fate and differentiation. *Proc Natl Acad Sci*, 111(28), pp.E2885–E2894. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24982138> [Accessed August 16, 2017].

Wang, S. et al., 2013. Human iPSC-derived oligodendrocyte progenitor cells can myelinate and rescue a mouse model of congenital hypomyelination. *Cell Stem Cell*, 12(2), pp.252–264. Available at: <http://dx.doi.org/10.1016/j.stem.2012.12.002> [Accessed August 16, 2017].

Wegner, M., 2008. A Matter of Identity: Transcriptional Control in Oligodendrocytes. *Journal of Molecular Neuroscience*, 35(1), pp.3–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18401762> [Accessed August 16, 2017].

Wernig, M. et al., 2008. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell stem cell*, 2(1), pp.10–2. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18371415> [Accessed August 16, 2017].

- Xie, H. et al., 2004. Stepwise reprogramming of B cells into macrophages. *Cell*, 117(5), pp.663–76. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/15163413> [Accessed August 16, 2017].
- Xie, K., Minkenberg, B. & Yang, Y., 2015. Boosting {CRISPR/Cas9} multiplex editing capability with the endogenous {tRNA-processing} system. *Proc Natl Acad Sci*, 112(11), pp.3570–3575.
- Xu, H. et al., 2015. Sequence determinants of improved CRISPR sgRNA design. *Genome Research*, 25(8), pp.1147–1157. Available at:
<http://genome.cshlp.org/lookup/doi/10.1101/gr.191452.115> [Accessed September 7, 2016].
- Xu, J. et al., 2015. Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell*, 16(2), pp.119–134. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/25658369> [Accessed August 16, 2017].
- Xu, L. et al., 2017. Empower multiplex cell and tissue-specific CRISPR-mediated gene manipulation with self-cleaving ribozymes and tRNA. *Nucleic Acids Research*, 45(5). Available at:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5389707/pdf/gkw1048.pdf>
[Accessed November 15, 2017].
- Yang, C. et al., 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*, 389(1), pp.52–65. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/17123746> [Accessed August 16, 2017].

- Yang, N. et al., 2013. Generation of oligodendroglial cells by direct lineage conversion. *Nature Biotechnology*, 31(5), pp.434–439. Available at: <http://www.nature.com/doifinder/10.1038/nbt.2564> [Accessed September 7, 2016].
- Yanger, K. et al., 2013. Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes & Development*, 27(7), pp.719–724. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23520387> [Accessed October 3, 2017].
- Yao, R. et al., 2014. Olig2 overexpression accelerates the differentiation of mouse embryonic stem cells into oligodendrocyte progenitor cells in vitro. *Dev Growth Differ*, 56(7), pp.511–517. Available at: <http://doi.wiley.com/10.1111/dgd.12150> [Accessed August 16, 2017].
- Zhang, Y. & Reinberg, D., 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & development*, 15(18), pp.2343–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11562345> [Accessed October 3, 2017].
- Zhao, X. et al., 2010. MicroRNA-Mediated Control of Oligodendrocyte Differentiation. *Neuron*, 65(5), pp.612–626. Available at: <http://dx.doi.org/10.1016/j.neuron.2010.02.018> [Accessed August 16, 2017].
- Zhao, X. et al., 2012. Specification and maintenance of oligodendrocyte precursor cells from neural progenitor cells: involvement of microRNA-7a. *Molecular biology of the cell*, 23(15), pp.2867–78. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22696677> [Accessed August 16, 2017].

Zhou, H. et al., 2018. In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR–dCas9-activator transgenic mice. *Nature Neuroscience*, pp.1-7.

Zopf, C.J. et al., 2013. Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression J. Kondev, ed. *PLoS Computational Biology*, 9(7), p.e1003161. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1003161> [Accessed August 16, 2017].

8 Appendix

8.1 Sequences of gRNAs used in this study

gRNA ID	Target gene	Target Sequence (5'-3')
g1	<i>Sox10</i>	TGGCATCTGGTGTGAGCATG
g2	<i>Sox10</i>	GTCTGGGAAGGGAGAGATAC
g3	<i>Sox10</i>	GGACCCTAAACCATTGCCAC
g4	<i>Sox10</i>	GTTGACAAAGGACTGGAGAG
g5	<i>Sox10</i>	GGAGAGTGGATAACTTGAAA
g6	<i>Sox10</i>	TGCGGTGCCTAGCACACAGT
g7	<i>Sox10</i>	AGGGAGGGAAGGGCAGAAGG
g8	<i>Sox10</i>	GGGGATGAAGGCAGCGCTGG
g9	<i>Sox10</i>	GGGCAGAGTGGAGGCGTCCT
g10	<i>Sox10</i>	GACGGGCCTAGGCAGCAGTG
g11	<i>Olig2</i>	GGTCGGCAGCGGGAACAGG
g12	<i>Olig2</i>	GCTAGCAAAAACCTCCAGGTG
g13	<i>Olig2</i>	GAATCTGCAAACTCGTCCGA
g14	<i>Olig2</i>	TGTAGCGTCAGGCCAGGGTA
g15	<i>Olig2</i>	GGGGCTGCTAAAGAAGTCTG
g16	<i>Olig2</i>	GGGCGAGCTGAAGAAGCGGG
g17	<i>Olig2</i>	GACCTGGGCAGGGAGAGGAA
g18	<i>Olig2</i>	GATTCGCGAGGAGGGATGGG
g19	<i>Olig2</i>	TGCTGAGAGGAGGCGTGTGC
g20	<i>Olig2</i>	ATGGCTGAGAATGAACACCG
g29	<i>Nkx6-2</i>	GGGGGCGCGCGCAGACGGGC
g30	<i>Nkx6-2</i>	GAAGGGCCCATTAGCGAGGC
g31	<i>Nkx6-2</i>	GGGCAAATTTACAGTCCCGG
g32	<i>Nkx6-2</i>	CGGGATTTATTTTCATCTCTA
g33	<i>Nkx6-2</i>	AGTTAGGCAAAGGTTTAAAT
g34	<i>Nkx6-2</i>	GCATACTGAGCTTGAGCTTT
g35	<i>Nkx6-2</i>	G TTCAGGGAAGAGCCTGGGA

g36	<i>Nkx6-2</i>	GGGCACTCCATCTTGGATCA
g37	<i>Nkx6-2</i>	AGCCCACACCACCACGAGCT
g38	<i>Nkx6-2</i>	GGAACCCCAGTACAGGGCTT
g39	<i>Nkx2-2 (promoter)</i>	GAGGGGGGAGGGGGAGAAAA
g40	<i>Nkx2-2 (promoter)</i>	GGAAAAGAGGGAGGGGGTCTG
g41	<i>Nkx2-2 (promoter)</i>	GAGAGAGGGGTGAAAAAAG
g42	<i>Nkx2-2 (promoter)</i>	GACATTAAAACCGCAAAGGT
g43	<i>Nkx2-2 (promoter)</i>	GGCGGGTCTTGGGAGTCAAG
g44	<i>Nkx2-2 (promoter)</i>	AGATGTGAAATTGTGGGTTT
g45	<i>Nkx2-2 (promoter)</i>	GCCCTCTAGAGCAAGATGAG
g46	<i>Nkx2-2 (promoter)</i>	GTCAATTAATTGCAAAGATG
g47	<i>Nkx2-2 (promoter)</i>	GTGTGGACTCGAGAGCGACG
g48	<i>Nkx2-2 (promoter)</i>	TAAGATGCAATTTGTTAAAA
g49	<i>Nkx2-2 (enhancer)</i>	TGTTAAGTGGGGCGAACAAG
g50	<i>Nkx2-2 (enhancer)</i>	GCTTAGCCTGGGGTGGGGGG
g51	<i>Nkx2-2 (enhancer)</i>	GTTCTGCTTCGCCTGGACTA
g52	<i>Nkx2-2 (enhancer)</i>	GCTCACAGCTGCTTGGGCCG
g53	<i>Nkx2-2 (enhancer)</i>	AGGCTCTAGAAACCCGGGGG
g54	<i>Nkx2-2 (enhancer)</i>	GTCACCCACGTCACATACTG
g55	<i>Nkx2-2 (enhancer)</i>	GCGTCCTGGGTGGTCGGACC
g56	<i>Nkx2-2 (enhancer)</i>	TGCTAAGCTGCGGACAATGA
g57	<i>Nkx2-2 (enhancer)</i>	GACAAATGTCCCGCTCCTGT
g58	<i>Nkx2-2 (enhancer)</i>	GAAGAGATAGTCCATTGAAA
g94	<i>Sox9</i>	TGCCTGGAAACTTCTGTGGG
g95	<i>Sox9</i>	CTGGAAGTCGGAGAGCCGAG
g96	<i>Sox9</i>	TTCCAAAATCCGGTCCAATC
g97	<i>Sox9</i>	CCCGGGTCCCCCGCTTGCCC
g147	<i>Sox9</i>	GGGTGACGAGACAGGAGGGG
g148	<i>Sox9</i>	TGGACCGGATTTTGGAAGGG
g149	<i>Sox9</i>	AGCTCTTGAGTCATCAGTGT
g150	<i>Sox9</i>	CGAGGTATCTAACGTGAAGG
g151	<i>Sox9</i>	GGCTCTCCGACTTCCAGCTC

g152	<i>Sox9</i>	GAACTGCCTGGAAACTTCTG
g153	<i>Sox10</i>	GGGCAGAAGGTGGAGCCTCC
g154	<i>Sox10</i>	TGGGTGTTGGGGATGAAGGC
g155	<i>Sox10</i>	GGGGGGCTGGGCAGAGTGGA
g156	<i>Sox10</i>	GAGACGGGCCTAGGCAGCAG
g157	<i>Sox10</i>	GTCTGATACCTACTGCAGGC
g159	<i>Nkx6-2</i>	CACTCCATGAAGGGCCCATT
g160	<i>Nkx6-2</i>	CAGCGTCCGGCCCGTCTGCG
g161	<i>Nkx6-2</i>	GCTGATTGGCTGAAGGCACT
g162	<i>Nkx6-2</i>	CCTTCATGGAGTGCATGGCC
g163	<i>Nkx6-2</i>	TGGAGTGCATGGCCGGCGGG
g175	<i>Olig2</i>	GCTGGGCCGGCCCGAGGGGC
g176	<i>Olig2</i>	TCGGGCCGGCCAGCTCATT
g177	<i>Olig2</i>	GGGCCCGGCGGGTTTCATTG
g178	<i>Olig2</i>	TGGCGGTGGTCCGGGCGGGG
g179	<i>Olig2</i>	TTACAGACCGAGCCAACACC
g180	<i>Olig2</i>	AGAGCCAGGTTCTCCTCCGC
g181	<i>OLIG2</i>	ACGGCGGCAGCGGTGGCG
g182	<i>OLIG2</i>	CCTGGCGAGTTTCATTGA
g183	<i>OLIG2</i>	CCACCACAGAGTCAGGTTG
g184	<i>Sox10</i> (for <i>Sox10</i> tagging)	CGACTCTATCCCGACCTTAG
g185	<i>Sox10</i> (for <i>Sox10</i> tagging)	ACTCTATCCCGACCTTAGAG

Table 8-1 | List of gRNAs used in this study and their sequences.

8.2 List of custom-made RNA-FISH probes and their sequences

Target gene	Probe #	Probe Sequence (5' to 3')
<i>Sox10</i>	1	tgataggtcttgttctcg
<i>Sox10</i>	2	ctgttccttcttgacctg
<i>Sox10</i>	3	cacggggaactgtcatcg
<i>Sox10</i>	4	tcatgggcctcttgacgtg
<i>Sox10</i>	5	ctgtgcccacaccatgaag

<i>Sox10</i>	6	tactggtcggctagctt
<i>Sox10</i>	7	tgagctcagcattgtggag
<i>Sox10</i>	8	agagcttgccctagtgctt
<i>Sox10</i>	9	ttgtcactttcgttcagca
<i>Sox10</i>	10	ctcagcctcctcaatgaag
<i>Sox10</i>	11	gatggtccttttgtgctg
<i>Sox10</i>	12	ccgaggttggtactttag
<i>Sox10</i>	13	tgagcctgaatagcagcag
<i>Sox10</i>	14	tgggttcccatctgacatg
<i>Sox10</i>	15	gactgcagctctgtctt
<i>Sox10</i>	16	cacgttgccgaagtcgatg
<i>Sox10</i>	17	ccatgttgacattacctc
<i>Sox10</i>	18	ccagctcagtcacatcaa
<i>Sox10</i>	19	cgttgggtggcaggtattg
<i>Sox10</i>	20	cgagtagctaccacatgg
<i>Sox10</i>	21	tggtggcttgagatccag
<i>Sox10</i>	22	cttggcatccacaccagg
<i>Sox10</i>	23	cgatctgggaagtggacgg
<i>Sox10</i>	24	gcagactgagggaagtgt
<i>Sox10</i>	25	atggaggggaaggcggagc
<i>Sox10</i>	26	tagtcaaactggggctgtg
<i>Sox10</i>	27	cctgagggctgatggtcag
<i>Sox10</i>	28	cctgcatggccataatagg
<i>Sox10</i>	29	gagatggcagtgtagagg
<i>Sox10</i>	30	tgtgtgggactgtggatt
<i>Sox10</i>	31	aaggtcgggatagagtcgt
dCas9	1	aagatctcttcagatagcaga
dCas9	2	ctctccagtctgtggaagaag
dCas9	3	gtcgaagttgctcttgaagttg
dCas9	4	gtcgtatctcttgatcatagag

dCas9	5	aaatctctttgtacttctcagg
dCas9	6	gtccttcaggaatgggtaaaaa
dCas9	7	gttcttatcgaagttggcatc
dCas9	8	tatacacggtgaagtactcgta
dCas9	9	tcacgtatttcactttggtcag
dCas9	10	tccttgataatttcagcagat
dCas9	11	cagcacgatacttccagaatg
dCas9	12	gatcatctctctgtcctcaaac
dCas9	13	tcacactttgtcgtcgaacag
dCas9	14	catgaagtttctgttggcgaag
dCas9	15	catttcgatcacgatgttctcg
dCas9	16	attctgcaggtagtacaggtac
dCas9	17	gatggagtcgtccttcagaaag
dCas9	18	tcagattgtcgaactttctctg
dCas9	19	ttgtcattctcgtcgacttag
dCas9	20	ttcaggggtgatcactttcactt
dCas9	21	gaaatccttccggaatcggac
dCas9	22	tgatctcgcgcactttgtaaaa
dCas9	23	cttagggtactttttgatcagg
dCas9	24	cttttcacgatattcacttgg
dCas9	25	caggatagactctttgctgaag
dCas9	26	tttcacactcttcagttcttg
dCas9	27	cagaaagtcgatgggattcttc
dCas9	28	tacttaggcagcttgatgatca
dCas9	29	ggtacaggaagttcacatattt
dCas9	30	aaacagctgtttctgctcatta
dCas9	31	caggatcactctcttgagaaac
dCas9	32	gacagcactttgtccagattag
dCas9	33	agggtaaacaggtggatgatat

Table 8-2 | List of custom-made RNA-FISH probes and their sequences.

8.3 List of key PCR primer used in this study

Position	Primer name	Sequence (5' to 3')	BsaI overhang	BsmBI overhang	BsmBI overhang for the position
1	gRNA_multiplex_1_F	AGCTAGGTCa ACTg gACGGCGTCt TA GGGATATGCTGTCaAGGTCGGCCAGGAAGAGGG	ACTG	TAGG	TAGG
	gRNA_multiplex_1_R	CGACTGGTCt TCG TCATCGTCt TC TCGATGACGGCATATCAAAAAAAGCACCGACTCGGT	TCGT	CTCG	CGAG
2	gRNA_multiplex_2_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA GAGATaAGGTCGGCAGGAAGAGGG	ACTG	CGAG	CGAG
	gRNA_multiplex_2_R	CGACTGGTCt TCG TCATCGTCt G CTGATGACGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	GCTG	CAGC
3	gRNA_multiplex_4_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA CAGCGAAGCGATTaAGGTCGGCAGGAAGAGGG	ACTG	CAGC	CAGC
	gRNA_multiplex_4_R	CGACTGGTCt TCG TCATCGTCt G CCATATGACGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	GCCT	AGGC
4	gRNA_multiplex_5_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA AGGCGAGCTAGTTaAGGTCGGCAGGAAGAGGG	ACTG	AGGC	AGGC
	gRNA_multiplex_2_R	CGACTGGTCt TCG TCATCGTCt TC TTGGCATGACGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	TGGC	GCCA
5	gRNA_multiplex_3_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA GATATGCTATCAaAGTCGGGCAAGGAAGAGGG	ACTG	GCCA	GCCA
	gRNA_multiplex_6_R	CGACTGGTCt TCG TCATCGTCt G CCATGACCGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	GCAC	GTGC
6	gRNA_multiplex_7_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA GTCCGAAGCGATTaAGGTCGGCAGGAAGAGGG	ACTG	GTGC	GTGC
	gRNA_multiplex_7_F	CGACTGGTCt TCG TCATCGTCt TC TCGATGACGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	TTCG	CGAA
7	gRNA_multiplex_8_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA GAAAGACTaAGGTCGGCAGGAAGAGGG	ACTG	CGAA	CGAA
	gRNA_multiplex_8_R	CGACTGGTCt TCG TCATCGTCt G TAGATGACGGATATTATCAAAAAAAGCACCGACTCGGT	TCGT	GTAG	CTAC
8	gRNA_multiplex_6_F	AGCTAGGTCa ACTg gAGCGTCGTCt CA CTACGATATGCTATCAaAGTCGGGCAAGGAAGAGGG	ACTG	CTAC	CTAC
	gRNA_multiplex_3_R	CGACTGGTCt TCG TCATCGTCt TC TCGTCATATCAAAAAAAGCACCGACTCGGT	TCGT	TCGT	ACGA

Table 8-3 | List of PCR primers used for construction of Ai1 vectors

PCR name	Sequence (5' to 3')	Used for
D1135V_F	CGGCGGCTTCgtgAGCCCCACCG	Site-specific mutagenesis
D1135V_R	TACTTCTTAGGGTCCCAGTCCTTCTTTC	Site-specific mutagenesis
G1218R_F	GGCCTCTGCCagaGAACTGCAGAAG	Site-specific mutagenesis
G1218R_R	AGCATTCTCTTCCGGCCG	Site-specific mutagenesis
R1335E_F	CGACCGGAAGgagTACACCAGCAC	Site-specific mutagenesis
R1335E_R	ATGGTGGTGTCAAAGTACTTG	Site-specific mutagenesis
T1337R_F	GAAGgagTACagaAGCACCAAAGAGGTG	Site-specific mutagenesis
T1337R_R	CGGTCGATGGTGGTGTCA	Site-specific mutagenesis
Sox10_F	GGAGATCAGCCACGAGGTAA	V5 genotyping
V5_R	ACCGAGGAGAGGGTTAGGGAT	V5 genotyping

Table 8-4 | List of PCR primers used for site-specific mutagenesis.